

University of Oxford

School of Anthropology and
Museum Ethnography



Student No: 459736
Christopher Kavanagh
Green Templeton College

Individual Pains & Social Gains: The Personal & Social Consequences of Collective Dysphoric Rituals

Coursework submitted for fulfilment of the requirements for the
Degree of DPhil Anthropology

Word count: 103,423 (excluding references and appendixes)

Supervisors:
Prof. Harvey Whitehouse

Institute of Cognitive &
Evolutionary Anthropology
Centre for Anthropology and Mind

I hereby certify that this is entirely my own work unless otherwise stated ☒

Abstract

Individual Pains & Social Gains:
The Personal & Social Consequences of Collective Dysphoric Rituals

Christopher Kavanagh
Green Templeton College
Submitted in Hilary Term 2015/2016

This thesis presents the results of a multi-method exploration of the effects of collective dysphoric rituals on self-identification, group affiliation, and prosocial behaviour. Findings are presented from a worldwide sample of martial artists, student participants in artificial ritual experiments, and observers and performers of Shinto firewalking festivals in Japan. The thesis tests recent predictions of the Modes of Religiosity theory in regards to the psychological processes that underpin shared dysphoric rituals and various costly signalling theories concerning the group orientated consequences of participating in extreme ritual events. The results from the studies raise questions with the broader generalisability of recent findings linking collective dysphoric rituals and inclusive self-identification and urge for a more nuanced appraisal of associations with prosocial behaviour. Furthermore, the role of subjective positive assessment of dysphoric experiences is shown to be a topic that has been unduly overlooked and preliminary evidence is provided for a potential relationship with identity fusion. Methodologically the thesis presents a series of novel artificial ritual studies that offer initial evidence in support of shared dysphoria's ability to enhance cooperation and promote positive ingroup association.

Abstract.....	2
List of Acronyms & Abbreviations.....	8
Acknowledgements.....	9
Central Aim of the Thesis.....	11
Core Research Questions.....	13
Structure of the Thesis.....	15
CHAPTER 1: THEORETICAL BACKGROUND.....	19
1.1 The Cognitive Science of Religion.....	19
1.2 Definitions of <i>Religion</i>	25
1.3 Definitions of <i>Ritual</i>	30
1.4 Ritual as a ‘Costly Signal’.....	39
1.5 The Divergent Modes of Religiosity Theory.....	45
1.6 Cognitively <i>Optimal</i> & Cognitively <i>Costly</i> Religion.....	60
1.7 Identity Fusion & Group Bonding.....	65
1.8 Affective Experiences & Group Bonding.....	81
1.9 Dysphoria & Social Bonding.....	89
1.10 Dysphoric Ritual Studies.....	105
1.11 Summary of the Current Thesis.....	125
CHAPTER 2: BRAZILIAN JIU JITSU SURVEY.....	130
2.1 Background of Brazilian Jiu Jitsu.....	130
2.2 Rationale for Study.....	134
2.3 Methods.....	138
Ethics.....	138
Recruitment Methods & Sample.....	138
Questionnaire Content.....	139
2.4 Results.....	142
Data Preparation.....	142
Hypothesis 1 (Belt Whipping = More Fusion, More Relational Bonds).....	151
Hypothesis 2 (More Dysphoric = Higher Identity Fusion).....	153
Hypothesis 3 (More Dysphoric = Less Frequent).....	159
Hypothesis 4a (Higher Fusion = More Costly Sacrifice).....	160
Hypothesis 4b (Mediation of effects by Relational bonds & Meaningfulness).....	170
Hypothesis 5a (More Dysphoric = Greater Costly Sacrifices).....	174
Hypothesis 5b (More Dysphoric = Stronger Relational Bonds).....	184
Hypothesis 6 (More Dysphoric = Higher Status & Solidarity).....	189
Hypothesis 7 (Dysphoric promotion experiences = Greater reflexivity).....	192
2.5 Discussion.....	198
Conclusion.....	215
CHAPTER 3: ARTIFICIAL RITUAL EXPERIMENTS.....	219
3.1 Background & Rationale.....	219
Study 1: ‘Wailing Infant’ Artificial Ritual V1 (UK).....	226
Hypotheses.....	228
3.2 Methods.....	229
Ethics.....	229
Recruitment & Sample.....	230
Pre-testing of Stimuli.....	231
Experimental Procedure.....	232
Materials.....	242
3.3 Results.....	243

Data Preparation	244
Manipulation Checks	245
Hypothesis 1 (Crying condition = Higher Fusion, More Relational Bonds).....	248
Hypothesis 2 (Crying Condition = More Cooperation)	250
Hypothesis 3a (Positive Affect = Higher Fusion).....	253
Hypothesis 3b (Positive affect = Relational Bonds)	253
Hypothesis 4 (Physiological Response = More Reflexivity).....	256
3.4 Discussion.....	258
Study 2: 'Wailing Infant' Artificial Ritual V2 (Japan)	267
Protocol Revisions	268
Hypotheses	274
3.5 Methods	274
Ethics.....	274
Recruitment & Sample.....	276
Pre-Testing of Stimuli.....	277
Experimental Procedure	280
3.6 Results	309
Data Preparation	309
Manipulation Checks	311
Hypothesis 1a (Dysphoric = More Fusion)	315
Hypothesis 1b (Dysphoric = More Relational Bonds).....	316
Hypothesis 2 (Dysphoric = Ingroup Preference)	317
Hypothesis 3a (Positive Affect = More Fusion).....	324
Hypothesis 3b (Positive affect = More Relational Bonds)	324
Hypothesis 4 (Dysphoric ritual experiences = Greater reflexivity).....	327
3.7 Discussion.....	331
Conclusions.....	337
CHAPTER 4: DYSPHORIC FIREWALKING RITUALS	341
4.1 Rituals in Japan.....	341
4.2 Rationale for Study.....	347
Field Study- Extreme Shinto Firewalking Festivals.....	352
Hypotheses	352
4.3 Methods	353
Ethics.....	353
Field sites.....	355
Recruitment & Interviews.....	358
Sample	359
Questionnaire Content.....	359
4.4 Results	363
Data Preparation	363
Hypothesis 1 (Post Event= More Prosocial).....	365
Hypothesis 2: (Post-Event = More Inclusive Social Identity)	369
Hypothesis 3 (Performers = More Fusion & Core Bonds with Relational Targets).....	372
Hypothesis 4 (Higher Fusion= More Costly Sacrifice)	376
Hypothesis 5 (More dysphoric experiences = Higher Fusion)	386
Hypothesis 6 (More pain = Prosocial behaviour from ingroup observers).....	389
4.5 Discussion.....	391
Conclusion.....	401
CHAPTER 5: CONCLUSIONS.....	408
General Conclusions	408
Avenues for Future Research	420
Concluding remarks.....	422

APPENDICES	425
APPENDIX I. BJJ Questionnaire	425
APPENDIX II. Wailing Infant Questionnaire V1	432
APPENDIX III. Wailing Infant V2 Video Instructions / Ritual Sequence	440
APPENDIX IV. Wailing Infant V2 Final Poster (Japanese).....	443
APPENDIX V. Wailing Infant V2 B-IAT Instructions	444
APPENDIX VI. Wailing Infant V2 Tajfel Allocation Tasks	445
APPENDIX VII. Shinto Firewalking Questionnaire	447

BIBLIOGRAPHY.....	454
--------------------------	------------

Index of Figures

Figure 1-1: Arousal level compared with frequency in 645 rituals.....	49
Figure 1-2: Heat map of (a) dysphoric and (b) euphoric arousal against frequency	53
Figure 1-3: Direction of influence among the variables of Modes theory.....	54
Figure 1-4: Verbal measure of ‘organizational identification’	69
Figure 1-5: Component structure of in-group identification.....	70
Figure 1-6: Verbal measure of ‘in-group identification’	71
Figure 1-7: Pictorial measure of identity fusion	74
Figure 1-8: Verbal measure of identity fusion.....	74
Figure 1-9: Correlations between rates of fusion and willingness to fight or die.....	75
Figure 1-10: The circumplex model of affect.....	84
Figure 1-11: Differences in bonding and pro-group behaviours over three experiments.....	92
Figure 1-12: Ex.2: Differential cooperation in a weakest link game	94
Figure 1-13: Donations increase with higher perceptions of pain amongst devotees	96
Figure 1-14: Self-punishment increases only in the presence of those treated unfairly.	100
Figure 1-15: Illustration of heart rate patterns during a firewalking ritual.....	107
Figure 1-16: Mean donations to temple by festival participants	108
Figure 1-17: Endorsement of exclusive (Hindu) vs. inclusive (Mauritian) identities	109
Figure 1-18: Self-reported fatigue in a collective firewalking ritual	111
Figure 1-19: Conceptual framework of extreme rituals as a social technology	113
Figure 1-20: Depth of reflection to a high-arousal artificial ritual	118
Figure 1-21: Change in depth of reflection in artificial ritual	119
Figure 2-1: A belt whipping gauntlet during a BJJ promotion event	132
Figure 2-2: Bruises & welts received during belt whipping gauntlets.....	133
Figure 2-3: Belt-whippings performed without jacket are rated more severe	147
Figure 2-4: Confirmatory factor analysis for Triandis IDV/COL scale	149
Figure 2-5: Subjective experience of belt-promotion events.....	152
Figure 2-6: Decrease in dysphoric ratings with increased frequency of promotions	160
Figure 2-7: Fusion score for different levels of donations to the group	169
Figure 2-8: Schematic representation of the two mediator model (Hayes 2013)	171
Figure 2-9: Relational bonds mediate fusion’s effect on donations of time.....	172
Figure 2-10: Relational bonds and meaningfulness mediate fusion’s effect on risking life.....	173
Figure 2-11: Mediation of positive experience effect on donations of time.....	177
Figure 2-12: Mediation of positive experience effect on willingness to endorse risking life	179
Figure 2-13: Positive experience and donations to the group	183

Figure 2-14: Fusion and identification mediate positive exp. and relational bonds	188
Figure 2-15: Ratings of reflexive accounts of promotional experiences.....	195
Figure 2-16: Positive ratings across all frequencies of belt promotions	201
Figure 3-1: HR pre- and post-ritual.....	246
Figure 3-2: Affective response to baby sounds stimulus.....	247
Figure 3-3: Most common fusion images endorsed in both conditions.....	249
Figure 3-4: Cooperation in the PGG between conditions.....	251
Figure 3-5: Wooden idols with schematic faces.....	272
Figure 3-6: Affective profile of infant vocalisation stimulus	279
Figure 3-7: Group identity card	287
Figure 3-8: Still image from instructional video	289
Figure 3-9: Ritual Material initial layout (euphoric condition)	291
Figure 3-10: Positions for ritual performance	295
Figure 3-11: Example Tajfel Matrices	303
Figure 3-12: Affective response to infant sounds stimulus.....	312
Figure 3-13: Differential Affective response to infant sounds stimulus.....	313
Figure 3-14: Implicit positive association with tribal names in both sessions	322
Figure 3-15: Ratings of reflexive accounts of promotional experiences.....	329
Figure 4-1: A Buddhist firewalking festival held in Miyajima	345
Figure 4-2: A Shinto ice-water immersion festival held in Tōkyō.....	346
Figure 4-3: <i>Tengu</i> performer passing through the flames	356
Figure 4-4: <i>Mikoshi</i> group passing through the flames.....	357
Figure 4-5: Willingness to risk injury in pre/post periods	367
Figure 4-6: Monetary donations in pre/post periods	368
Figure 4-7: Inclusive/Exclusive identity endorsement in pre/post periods.....	370
Figure 4-8: Fusion & Bonds in pre-post periods for non-local observers	373
Figure 4-9: Fusion & Bonds in pre-post periods for local observers	374
Figure 4-10: Fusion & Bonds in pre-post periods for performers.....	375
Figure 4-11: Core bonds mediate fusion's effect on willingness to risk injury	383
Figure 4-12: Perceived pain by observers and performers.....	390
Figure 4-13: Identity endorsement post-ritual in Japan and Mauritius	392

Index of Tables

Table 1-1: Features of doctrinal & imagistic 'modes' of religiosity.....	48
Table 1-2: Predictions of the mnemonic effects of dysphoria in rituals	60
Table 1-3: The benefits of pain and associated processes.....	91
Table 1-4: Difference in contributions between effortful and easy tasks	116
Table 2-1: BJJ belt promotions questionnaire content	140
Table 2-2: Rotated factor loadings for response to belt promotion events	145
Table 2-3: Spearman's correlations for objective measures of severity	146
Table 2-4: Rotated factor loadings for Triandis IDV/COL scale	150
Table 2-5: Descriptive statistics for H1 variables	152
Table 2-6: Spearman's correlation matrix for dys./eup. experiences and group affiliation	154
Table 2-7: Regression models predicting identity fusion	158
Table 2-8: Descriptive statistics for self-reported sacrifice measures & donations.....	161

Table 2-9: Correlations for sacrifice and group affiliation measures	161
Table 2-10: Exploratory regression with willingness to donate time as outcome	163
Table 2-11: Hierarchical regression on willingness to donate time	164
Table 2-12: Exploratory regression with willingness to risk life as outcome	165
Table 2-13: Hierarchical multiple regression on willingness to risk life	166
Table 2-14: Correlation with donation measure	167
Table 2-15: Hierarchical ordinal regression on willingness to donate	169
Table 2-16: Spearman's correlation matrix for sacrifice measures, fusion and mediators	171
Table 2-17: Spearman's correlations for sacrifice measures and	174
Table 2-18: Hierarchical multiple regression on willingness to donate time	176
Table 2-19: Hierarchical multiple regression on willingness to risk life	178
Table 2-20: Hierarchical ordinal regression on willingness to donate	181
Table 2-21: Descriptive statistics for relational measures	185
Table 2-22: Exploratory regression of relational bond and sociometric knowledge	186
Table 2-23: Regression Models predicting relational bonds	187
Table 2-24: Spearman's correlations for group status and entitativity measures	189
Table 2-25: Regression models predicting group status and entitativity	191
Table 2-26: Reliability of ratings	193
Table 2-27: Hierarchical multiple regression on importance rating	197
Table 3-1: Instructions for Ritual Sequence	235
Table 3-2: Descriptive statistics for identity fusion and relational bonds	249
Table 3-3: Descriptive statistics for cooperation	250
Table 3-4: Spearman's correlations with cooperation in PGG	252
Table 3-5: Descriptive statistics for PANAS measures	253
Table 3-6: Hierarchical multiple regression on relational bonds S1	254
Table 3-7: Hierarchical multiple regression on relational bonds S2	255
Table 3-8: Descriptive statistics for PANAS measures	256
Table 3-9: Spearman's correlations for reflexivity and physio/affect measures	257
Table 3-10: Summary of Revised Experimental Conditions	273
Table 3-11: Tribal Names & Roles in first ritual task	287
Table 3-12: Instructions for Main Ritual Sequence (Stage 2)	294
Table 3-13: IAT targets	298
Table 3-14: Summary of Methodological Improvements	308
Table 3-15: Descriptive statistics for group fusion and identification measures	315
Table 3-16: Descriptive statistics for relational bonds	316
Table 3-17: Spearman's correlation matrix for identity fusion and relational bonds	317
Table 3-18: Descriptive statistics for IAT outcomes	321
Table 3-19: Descriptive statistics for PANAS measures (post-ritual)	324
Table 3-20: Pearson's correlations for relational bonds	324
Table 3-21: Hierarchical multiple regression on relational bonds S1	325
Table 3-22: Hierarchical multiple regression on relational bonds S2	326
Table 3-23: Reliability of SER Ratings	328
Table 3-24: Spearman's correlations for reflexivity and physio./affect. measures	330
Table 3-25: Multiple regression on reflexivity S2	330
Table 4-1: Shinto firewalking questionnaire content	362
Table 4-2: Rotated factor loadings for response to festival items	364
Table 4-3: Descriptive statistics for H1 variables	366

Table 4-4: Descriptive statistics for H2 variables	369
Table 4-5: Descriptive statistics for H3 variables	372
Table 4-6: Spearman's correlation matrix for fusion and sacrifice measures	376
Table 4-7: Multiple regression on costly sacrifice measures	377
Table 4-8: Exploratory regression with willingness to donate time as outcome	379
Table 4-9: Hierarchical regression on willingness to donate time	380
Table 4-10: Exploratory regression with willingness to risk injury as outcome	381
Table 4-11: Hierarchical regression on willingness to risk injury	382
Table 4-12: Exploratory regression with voluntary donation of prize	384
Table 4-13: Hierarchical regression on costly donation.....	385
Table 4-14: Spearman's correlation for firewalking past experience and fusion	386
Table 4-15: Hierarchical multiple regression on fusion with <i>mikoshi</i> performers.....	387
Table 4-16: Hierarchical multiple regression on fusion with local community	388
Table 4-17: Spearman's correlation matrix for pain and prosocial measures.....	389
Table 5-1: Summary of Findings for Core Q1.....	411
Table 5-2: Summary of Findings for Core Q2.....	414
Table 5-3: Summary of Findings for Core Q3.....	415
Table 5-4: Summary of Findings for Core Q4.....	417

List of Acronyms & Abbreviations

CSR = Cognitive Science of Religion
 MCI= Minimally Counterintuitive
 CREDs= Credibility Enhancing Displays
 IAT= Implicit Association Test
 BIAT= Brief Implicit Association Test
 SER= Spontaneous Exegetical Reflection
 MGP= Minimal Group Paradigm
 PGG= Public Goods Game
 MJP= Maximum Joint Profit (Tajfel Matrix allocation strategy)
 MD= Maximum Difference (Tajfel Matrix allocation strategy)
 MIP= Maximum Ingroup Profit (Tajfel Matrix allocation strategy)
 Modes Theory= The Divergent Modes of Religiosity Theory

Acknowledgements

The limitations in this thesis are all my own responsibility but its successes owe a great debt to the many people who have assisted me over the past four years. First, I need to express my gratitude to my supervisor, Prof. Harvey Whitehouse, it was his theories that inspired me to pursue my research and which continue to motivate me to improve my own ideas. Harvey has been a continued source of inspiration, as well as a patient and kind advisor. I have learnt a lot while under his supervision and I appreciate all the trust he has extended to me during my fieldwork in Japan. I look forward to continuing our collaborations long into the future. Special recognition also needs to be extended to Dr. Jonathan Jong, my boss in the ICEA lab and the *Ritual, Community, & Conflict* project. He has spent countless, unrewarded hours providing advice and encouragement, chastising when necessary, and most of all, being a friend. Without his council, insight, and 'pastoral support' I am not sure if I would have made it this far.

I also owe a great deal of gratitude to Prof. Masaki Yuki in Hokkaido University. A number of years ago he opened his lab to me and has now kept the door open long past our original plan for a six-month visit. Over the past two years in almost every respect he has served as a second unofficial supervisor, engaging in critical discussions about theories and data, and recommending me new theories and approaches. I have been welcomed as a member of his lab at the Culture, Social Ecology, and Psychology Lab in Hokkaido and it now feels like my 2nd home.

In mentioning the lab, it would also be remiss if I didn't also give special mention to Shuhei Tsuchida, a constant companion in all of my field research endeavours and an invaluable collaborator during experiments. Rob Thompson, Yo Nakawake, Wataru Toyokawa, and Masaki Suyama, also provided invaluable assistance and stimulating friendship. My gratitude

also extends to my long suffering friends, Doug McNaught and Gerard Reid, both of whom helped me to push through the final few months and even provided some last minute proofreading to boot.

I also wish to acknowledge the contribution made by all the various officials and volunteers of the *Bikuni* Shrine in Shakotan, *Kotohiro* Shrine in Furubira, *Daishoin* Temple in Miyajima, *Teppozu Inari* Shrine in Tokyo, and *Fukugonji* Temple in Nagoya. Without their willingness to open their communities to a strange foreign researcher during their busiest months little of my work in Japan would have been possible.

Ultimately, I reserve my deepest thanks for my wife- Ran, my son- Kai, and my parents- Irene and Michael, back in Ireland. Ran has endured a lot throughout the past four years of my research and writing, and it would have been impossible to complete this thesis without her unrelenting care and support. Any expression of gratitude would be insufficient to recognise all she has done. Similarly, my parents have been a constant well of emotional support and encouragement throughout my DPhil and were particularly helpful in keeping me motivated to make the final push. Finally, I would like to thank my son Kai for his incessant and welcome attempts to distract me with talk of dinosaurs.

Funding for this research was provided by two sources: the ESRC *Ritual, Community, & Conflict* project and a large grant from the John Templeton Foundation as part of the *Religion's Impact on Human Life* Project. The funds supplied covered the costs of materials, field research expenses, recruiting participants, and financial support for research assistants.

Central Aim of the Thesis

The research presented in this thesis is intended to contribute to the Cognitive Science of Religion (CSR), a rapidly developing field that employs quantitative and experimental methods to analyse beliefs and practices commonly associated with religion and religious traditions (J. L. Barrett, 2007, 2011; Pyysiäinen, 2013). Specifically, this thesis seeks to provide a detailed examination of, and novel evidence concerning, the individual psychological impact and wider social effects of dysphoric rituals.

Dysphoric rituals have been found to be cross culturally prevalent (Atkinson & Whitehouse, 2011) and are central to a number of influential theories addressing the persistence and popularity of religions and ritual activity worldwide, including costly signalling theories (Henrich, 2009; Sosis & Alcorta, 2003; Sosis, 2005) and the 'Divergent Modes of Religiosity' theory (Whitehouse, 2000, 2004). The conflicting predictions and the areas of compatibility that these theories present for dysphoric rituals is outlined in the literature review (Chapter 1) and explored through the novel research presented in Chapters 2, 3 & 4.

The new research reported in this thesis is drawn from a combination of laboratory-based experiments, online surveys, and field experimentation. This diverse combination of methods inevitably generates problems with comparison, however it also helps to address concerns of reliability that afflict findings derived from a single methodology. Results which can be replicated across methodologies increase confidence in the validity of the findings, whereas findings that are incongruent serve to problematize generalisations.

The samples employed in this thesis include a worldwide sample of Brazilian Jiu Jitsu practitioners, student participants from Oxford (UK) and Hokkaido (Japan), and performers and observers of extreme firewalking ritual events conducted at Shinto shrines in Northern Japan. This diverse sample was chosen due to the author's substantial experience with Japanese religious groups and to avoid common criticisms levelled at the CSR field of an overreliance on participants who are from Christian cultures and a 'Western Educated Industrialised Rich and Democratic' (WEIRD) background (Henrich, Heine, & Norenzayan, 2010). Furthermore, most treatments of ritual to date have focused on communities with significant levels of religious belief, whereas Japan despite high levels of participation in religious ritual events (Reader & Tanabe, 1998; Reader, 1990) is purportedly the second least religious country in the world (WIN-Gallup International, 2012; WinGallup, 2014). This unusual combination means that examining ritual activity in Japan provides a unique opportunity to address the impact of rituals and ritualised behaviour in a cultural context where overt devotion to a particular religious group or doctrine is largely absent.

Overall, the central theoretical aim of the dissertation is to provide novel insight into the individual and social effects of dysphoric ritual events through a series of innovative field and lab based studies. Methodologically the aim of the thesis is first, to provide an account and critical analysis of the novel research methodologies developed for this thesis and second, to demonstrate the utility of employing multiple methodologies across different contexts in order to gain broader insight. Finally, although the primary evidence presented

in this thesis is quantitative and relies on statistical analysis, each chapter also seeks to provide relevant qualitative contextual and ethnographic information.

Core Research Questions

In order to ensure that the central aims of this thesis are achievable it is important to establish clear research questions. To be effective these should accurately represent predictions derived from the existing research literature and also be stated such that new insights are gained by answering them. With these goals in mind four core research questions were developed and are listed below along with a short elaboration of significant sub-questions. None of the studies reported in the thesis is intended to independently provide a definitive answer to all four questions but collectively the studies are intended to provide new evidence that adds to our existing knowledge.

Q1. Do collective dysphoric rituals promote affiliative bonds, and if so, do effects differ amongst participants and observers?

Specifically, do dysphoric rituals promote the intense relational bonds characteristic of identity fusion (Swann, Jetten, Gómez, Whitehouse, & Bastian, 2012; Whitehouse & Lanman, 2014) or do they invoke depersonalisation (Rosenberg, 1987) and the development of collective categorical ties characteristic of group identification (Hogg, 2006; J. C. Turner & Reynolds, 2010)? Are affiliative effects restricted to participants or is there a collective response that is also shared by observers (Fischer et al., 2014; Xygalatas et al., 2013a, 2013b). Are these bonds collective or directed at specific individuals (Yuki, 2011)? And do perceptions of familial ties (Swann & Buhrmester, 2015; Swann et al., 2014) or meaningfulness (Olivola & Shafir, 2013; Whitehouse & Lanman, 2014; Whitehouse, 2002c)

mediate effects? Does the performance of dysphoric rituals serve as a 'costly signal' of commitment (Bulbulia & Sosis, 2011; Henrich, 2009) to the group and lead to greater perceived status or higher levels of trust from other in-group members who observe the ritual?

Q2. Does participation in, or observation of, dysphoric rituals generate prosocial behaviour directed towards the ingroup?

Is there support for the preliminary findings from recent studies that those who engage in (Bastian, Jetten, & Ferris, 2014), or simply observe (Xygalatas et al., 2013a, 2013b), collective dysphoric practices will behave in a more prosocial manner towards their ingroup? And if so can this prosocial drive motivate a willingness to endorse or perform costly sacrifices for the group (Swann, Gómez, Huici, Morales, & Hixon, 2010; Swann et al., 2014)? Does the level of pain endured predict performers prosocial behaviour (Fischer & Xygalatas, 2014; Xygalatas et al., 2013b) or alternatively does it make them feel they have 'paid their dues' (van Bunderen & Bastian, 2014) and reduce their contributions? Does the perception of greater pain lead to compensatory pro-group efforts being made by other observing ingroup members?

Q3. Are the individual and social effects generated from collective dysphoric rituals distinct from those of collective, high-arousal euphoric rituals?

Is there a distinctive effect from collective dysphoric arousal that is not observed with similar levels of collective euphoric arousal (Y. I. Russell, Dunbar, & Gobet, 2011; Xygalatas, 2014)? And what is the role of euphoric affective response to dysphoric ritual events (Fischer et al., 2014)? Does the priming of dysphoric experiences produce affiliative and

prosocial outcomes that can be distinguished from those produced by other collective group experiences that lack a dysphoric element?

Q4. How do collective dysphoric rituals identities impact self-identity?

Do they lead to the endorsement of more inclusive personal identities, as found recently with Hindu piercing rituals in Mauritius (Xygalatas et al., 2013b)? Or do they promote more exclusive identities and potentially outgroup discrimination, as suggested by cross cultural correlations observed between ritual attendance and endorsement of suicide attacks (Ginges, Hansen, & Norenzayan, 2009) and existing research on identity markers and intergroup discrimination (Diehl, 1990; Tajfel, Billig, Bundy, & Flament, 1971a; Tajfel & Turner, 1985; Yoshida & Kubota, 1994)?

These questions are intended to be empirically tractable and will be examined through data collected from a worldwide survey of martial artists (Chapter 2), participants in controlled laboratory experiments employing an artificial ritual protocol (Chapter 3), and a quasi-experimental field survey conducted with participants at Shinto firewalking festivals in Japan (Chapter 4).

Structure of the Thesis

To provide a coherent structure to this multi-faceted examination of collective dysphoric ritual events this thesis is organised into the following five chapters:

Chapter 1: Literature Review

The first chapter provides a critical literature review and seeks to establish the current state of the field in research on dysphoric rituals and ritual practices more generally. The core approach of the cognitive science of religion and the advantages it offers are briefly

summarised, along with important debates on the utility of the concepts 'religion' and 'ritual'. The operational definitions for both concepts employed throughout the thesis are also provided and justified. However, the primary focus of the chapter is to explore the major theoretical accounts in cognitive anthropology that address dysphoric rituals. As a result, the divergent modes of religiosity theory and costly signalling theories are evaluated critically and a review is provided of recent relevant empirical studies (Bastian, Jetten, & Ferris, 2014; Bastian, Jetten, Hornsey, & Leknes, 2014; Xygalatas et al., 2013b). The related key psychological theories of identity fusion and group identification are also discussed. Finally, the chapter concludes by summarising how the literature reviewed will be addressed through the novel research presented in the thesis.

Chapter 2: Dysphoric promotion rituals in Brazilian Jiu Jitsu Schools

The second chapter details the results from an online questionnaire conducted with a worldwide sample of Brazilian Jiu Jitsu practitioners. The questionnaire was designed to explore the relationship between an individual's experiences of belt promotion practices and whether this influences their relationship with their school and their prosocial tendencies. Specifically, the chapter addresses whether dysphoric belt whipping 'gauntlets' accord with predictions of Modes theory and are associated with less frequent promotions, stronger relational bonds between members, and generate identity fusion with the in-group. In the results and discussion section the significance of positive and negative *subjective* assessments of dysphoric events are discussed.

Chapter 3: Artificial Ritual Studies

The third chapter reports the results of two experiments conducted in the UK and Japan that set out to develop a novel lab-based ‘artificial ritual’ methodology. The protocol was designed to model the core features of collective ritual practices while varying their affective profile between conditions. The first study, conducted with volunteers in Oxford, explored whether shared dysphoric rituals promote cooperation. To achieve this a between-groups study was conducted in which an artificial collective ritual was performed by small groups of volunteers with a dysphoric audio stimulus in one condition and a comparable neutral stimulus in the other. The second study, conducted in Japan, revised the artificial ritual methodology substantially; increasing the interactive elements of the collective ritual, improving the quality of the variable affective stimuli, and adding a euphoric condition to enable distinctions to be made between general high arousal effects and specific responses to dysphoric elements. Intergroup allocation tasks using Tajfel matrices were also employed in order to permit a more fine-grained examination of ingroup preference, along with an implicit association task (IAT) to examine implicit group preferences. The chapter concludes with a discussion of how the findings from both experimental studies compare and how they relate to the broader research literature.

Chapter 4: Japanese Fire walking festivals

Chapter four is based on questionnaire data collected at two extreme fire walking festivals performed at Shinto shrines in two small fishing towns in Northern Japan and seeks to replicate recently reported findings from dysphoric rituals in Mauritius (Xygalatas, et al., 2013b). The chapter is organised around reporting the quasi-experimental field study

but begins with a general summary of the religious and ritual context of Japan. Following this, the demographic and social context of the field sites visited are described and details are provided of the methodology employed. Results from questionnaires collected from performers and observers in pre and post firewalking periods are then reported, analysed, and discussed. Primarily this chapter addresses the question of whether different responses are generated in performers and observers of extreme rituals: examining reported identity fusion with a variety of targets, the endorsement of inclusive/exclusive personal identities, and the extent of prosocial attitudes and behaviour.

Chapter 5: Conclusion

Chapter five summarises the findings from the previous chapters and highlights the areas of continuity and disagreement between the different methodologies and field sites. How the findings contribute to the existing research literature is also explored, along with a discussion of the core issues involved with attempting to disentangle dysphoric practices from their euphoric interpretations. Finally, a summary of potential avenues for future research and some general concluding remarks are provided.

CHAPTER 1: THEORETICAL BACKGROUND

1.1 The Cognitive Science of Religion

This project assumes an approach and methodology consistent with the Cognitive Science of Religion (CSR); an interdisciplinary initiative focused on how religions and religious concepts worldwide, despite their apparent diversity, are shaped and constrained by universal features of human cognition (J. L. Barrett, 2007). Therefore, before addressing the specific theories from this field that this thesis seeks to explore, it is necessary to briefly summarise what this approach entails; in terms of its characteristic features, methodological assumptions, and some criticisms it faces.

The field of CSR, in relative terms, is a newcomer to the academic scene and developed in large part from the impact of the 'cognitive revolution' (G. A. Miller, 2003) on the social sciences. The cognitive revolution impacted certain fields in the anthropology discipline dramatically (Blount, 2011), but it had influenced only a small number of anthropological and psychological examinations of religion up to the 1980s (Guthrie, 1980; Spiro, 1982). However, during the 1990s many cognitive and evolutionary studies of religion rapidly appeared following the publication of several influential theoretical works (J. L. Barrett & Keil, 1996; Boyer, 1994; Guthrie, 1993; Lawson & McCauley, 1990; Rappaport, 1999; Whitehouse, 1995). During this period, empirically minded anthropologists drew inspiration for new theories on religion from a new raft of *cognitive* approaches to culture, including Dan Sperber's epidemiological examination of cultural transmission (1975, 1996)

and Richerson & Boyd's (1976, 1985, 2005) 'gene culture coevolution' or 'dual inheritance' theory.

The impact of these works was substantial and pioneering CSR researchers quickly demonstrated how such approaches could be applied to the topic of religion. Focusing on rituals, for instance, Whitehouse (1992, 1995) proposed that the dramatic differences between different types of ritual performances were related to the differential activation of semantic and episodic memory systems. Boyer (1994) and Barrett (1998) meanwhile argued that the majority of supernatural beings found in religious systems could be categorised according to a core *minimally counterintuitive* (MCI) typology. This meant that supernatural agents the world over, despite their seeming disparity, shared a core template that grabbed attention due to the presence of a *limited* amount of features that violated ontological assumptions¹.

CSR researchers were not the first to advocate evolutionarily or cognitively informed approaches to examining culture. In anthropology in particular there have been a number of influential figures who promoted cognitively informed research. Claude Lévi-Strauss (1963, 1971, 1983), for example, based his entire career around efforts to uncover how culture was determined by universal concepts of cognition. Furthermore, during the 1960s, inspired by the cognitive revolution and Chomskian linguistics (Chomsky, 1965), a specific field labelled *cognitive anthropology* emerged in the US. This field still exists today and has

¹ These are just two illustrative examples, both of which have since faced significant challenges (Harmon-Vukić, Upal, & Sheehan, 2012; McCauley & Lawson, 2002; McCauley, 2001; Purzycki & Willard, 2015; Upal, 2010; Wiebe, 2004) and undergone substantial revisions (J. L. Barrett, 2008, 2012; Gregory & Barrett, 2009; Whitehouse & Lanman, 2014; Whitehouse, 2004).

amassed a substantial literature dedicated to addressing the universality of classificatory taxonomies, social systems (particularly kinship), and perceptual abilities (Conklin, 1962; Frake, 1969).

Although there is influence in CSR from such pioneering efforts, CSR researchers today are cautious in their assessments of its relevance (Boyer, 2013). In particular, they note that a number of the earlier influential cognitive approaches, such as Lévi-Strauss' binary analyses, suffered from a lack of appropriate empirical data analysis, the misapplication of linear progressive models of evolution, and a reliance on models of cognition that have been largely discredited by modern research. Consequently, the primary influence from a lot of earlier cognitive and evolutionary anthropological efforts is acknowledged to be inspirational in nature rather than in providing useful theoretical models to test (Boyer, 2013; Guthrie, 2013)².

Partly due to the awareness of earlier pitfalls, a defining characteristic of the modern CSR approach is that it seeks to be an explicitly interdisciplinary scientific endeavour and is expected to be updated to accord with the best supported findings from relevant fields of evolutionary and cognitive research. This does not preclude efforts to produce original research but rather incentivises researchers to focus on gathering empirical data that can be used to test the predictions of both existing and novel theoretical models. The ultimate goal

² There are important exceptions, with Durkheim's theories currently experiencing something of a resurgence (Shariff, Purzycki, & Sosis, 2014; Whitehouse, 2013a; Xygalatas et al., 2011) and the insights derived from the 'cognitive anthropologists' explorations of folk taxonomies, kinship structures, and religious systems still recognised (Spiro & D'Andrade, 1958; Spiro, 1965, 1982).

of CSR is to build a reliable, cumulative body of knowledge in regards to the topic of religion (J. L. Barrett, 2007; Cohen, Lanman, Whitehouse, & McCauley, 2008).

Accordingly, in terms of both methodology and goals, CSR stands in stark contrast with the postmodern and deconstructivist approaches that have dominated mainstream social and cultural anthropology since at least the 1980s (Clifford & Marcus, 1986; Geertz, 1973). In contrast to the cumulative efforts just described these approaches are characterised by a tendency to dismiss the theories of previous generations and ideological opponent as being not only misguided but morally compromised.³

Furthermore, although postcolonial, deconstructivist, and postmodern approaches to religion are diverse in many respects, they tend to be united by their extreme scepticism, if not outright hostility, towards empiricism or the utility of scientific methodologies for understanding any aspect of religion, or human culture more generally (Geertz, 1973). Deconstructing analytical concepts and theories (Derrida, 1997) has been prioritised instead over any attempt to build predictive and testable models, due to an emphasis on reflexivity and a desire to avoid ethnocentrism (Clifford & Marcus, 1986). Such approaches have repeatedly been applied to criticise the very validity of the concept of religion (Asad, 1993; Fitzgerald, 1999; Harvey, 2013).

Consequently the empirical efforts in CSR research to identify and account for cross culturally recurrent features of religion and religious behaviour are viewed as offering little

³ See for instance, the infamous personal clash between Gananath Obeyesekere and Marshall Sahlins over Captain Cook's death (Borofsky, 1997; Obeyesekere, 1997; Sahlins, 1996); Derek Freeman's critique of Margaret Mead's research in Samoa (Freeman, 1983; Mead, 1943; Shankman, 2009) and the extreme dismissals of both Jared Diamond and Napoleon Chagnon's work (Correia, 2013; Tierney, 2001).

more than a renewal of discredited sociobiological theories (Van Slyke, 2011) that promote an imperialistic, 'vicious' reductionism (Polkinghorne 2001; Cho & Richard K. Squier 2008a; Cho & Richard K. Squier 2008b: 412) and take a decontextualized, essentialised 'Other' as its subject. Critics thus charge that the CSR approach offers simplistic, deterministic accounts that are blind to many of the socio-historical factors that influence religious phenomena (Peterson, 2006; Stroeken, 2011) and that they are too committed to speculative assumptions about cognitive modularity (Watts 2014).

The critics are right to condemn the dismissal of socio-historical factors and to caution against the uncritical acceptance of popular psychological and cognitive theories. However, it is disputable whether such insights are really ignored by CSR researchers as, setting aside provocative book titles, the field as a whole unequivocally eschews any attempt to present any single explanations for the complex cultural phenomena ascribed the label 'religion' (Bulbulia et al., 2013; Cohen et al., 2008; McCauley, 2013; Slingerland & Bulbulia, 2011). Instead, there has been a deliberate trend within the CSR field to advocate for fractionated the topic of religion "into numerous different traits, each of which must be explained on its own account" (Whitehouse, 2008: 35). Hence, Barrett (2007) has even argued that approaching religion in an incremental, piecemeal fashion is a defining feature of the CSR approach. Additionally, the importance of recognising an interactive influence from an individual's surrounding social and ecological environment and their inherited biological makeup has long been promoted by CSR researchers (Rappaport, 1999;

Whitehouse & Laidlaw, 2004; Whitehouse, 2004) and is currently being emphasised with renewed vigour (Bulbulia et al., 2013; Whitehouse, 2008).⁴

CSR remains a relatively immature field but possessing a unifying empirical framework has enabled the field to progress rapidly; with major theoretical works continuing to appear regularly and older theoretical models being reviewed, revised, and challenged (Baumard & Boyer, 2013; Gregory, forthcoming; Norenzayan, Henrich, & Slingerland, 2013; Purzycki & Willard, 2015; Purzycki et al., 2016; Shariff, Willard, Andersen, & Norenzayan, 2015; Upal, 2010; Whitehouse & Lanman, 2014). Therefore, despite the short 25 year history a substantial body of empirical research has already been amassed and the sophistication of experimental methods displays a clear upward trend (Atkinson & Whitehouse, 2011; Atran & Ginges, 2012; K. Banerjee, Haque, & Spelke, 2013; Bulbulia et al., 2013; Cohen, Ejsmond-Frey, Knight, & Dunbar, 2010; Turchin, Whitehouse, Francois, Slingerland, & Collard, 2012; J. A. Watts, Sheehan, Atkinson, Bulbulia, & Gray, 2016; Xygalatas, Konvalinka, Bulbulia, & Roepstorff, 2011).

Yet, any balanced assessment of the existing literature must also recognise that the CSR field is currently theory laden with significant empirical gaps (J. L. Barrett, 2005, 2011; Jong, Kavanagh, & Visala, 2015). Hence, there is now a clear need for consistent efforts to be made to put the predictions of existing theories to the test. Fortunately, in recent years, there have been efforts made to do just that, with various large-scale collaborative projects

⁴ Purzycki, Haque & Sosis (2014: 85), for example, state that any viable account of religion must give “equal attention to the nature of the human mind, the natural and social environment, history, and all of the forces which exert influence upon them and their relationships” and “explore not only how religious systems persist, but also how they respond to environmental changes”.

primarily focused on theory testing, such as the ESRC funded ‘Explaining Religion’ project (2007-2010), and the on-going ‘Ritual, Community and Conflict’ project (2011-2016). The research presented in this thesis was partly funded by the latter project and seeks to contribute to theory testing efforts by addressing outstanding issues surrounding dysphoric rituals and their effects on individual cognition and group bonding. The relevant ritual theories from the CSR and their associated predictions are surveyed in greater detail later in this chapter.

1.2 Definitions of *Religion*

Clarity with definitions is important as it helps to delineate precisely what is under discussion and avoid unnecessary semantic disputes. Yet, some degree of semantic debate is unavoidable in academic research and as a result to differing theoretical and methodological agendas definitional debates are unfortunately an unresolvable preoccupation for researchers (Franek, 2015).

The CSR field advocates fractionating the diffuse category of *religion* into tractable components for research purposes and is ultimately agnostic to the existence of a fundamental, universal category of religion but some individual researchers do advocate particular positions on the topic. Boyer (2011) and Bloch (2008, 2014), for example, state categorically that there really is no such thing as religion⁵ and that “the separation of religion from the transcendental social in general is, even in the places where it appears at first to exist, superficial and transient” (2008: 2060). Alternatively, Norenzayan (2009)

⁵ A view somewhat incongruous with the titles and claims made in Boyer’s previous works, particularly *Religion Explained* (1994).

argues that religion is a ‘human universal’ and Bulbulia (2005: 72-73) similarly outlines “that there is one human religion with variation at the margins” and this represents: “a panhuman capacity... based on a psychological architecture that ... is largely invariant across culture and era [and] reliably produces functionally similar expressions and commitments, irrespective of labelling and convention”.⁶

Clearly then, although there is a general consensus in the field that no single explanation of religious phenomenon is possible or sufficient to account for the diversity of religious phenomenon across the globe (J. L. Barrett, 2007, 2011; Cohen et al., 2008), disagreements surrounding the definition and usefulness of the category *religion* remain persistent (Franek, 2015). In the face of such debates it seems important to begin this thesis by explicitly addressing how the term *religion* will be operationalised in the following chapters and offering a short justification for the definition selected.

There is merit to the arguments that the core reference of the category religion as it is commonly understood is limited to religious institutions in the modern world and that it is thus an artificial and ethnocentric term (Asad, 1993; Fitzgerald, 1999; Harvey, 2013; McCutcheon, 1997) that can be misleading when applied to cultures outside the Abrahamic-Judaeo sphere (Fitzgerald, 2003, 2004; Josephson, 2011). However, such criticisms tend to overlook two important counter arguments. The first is that the same level of scepticism and accusations of artificiality can be applied to practically any term that can be used to discuss aspects of human societies cross-culturally. Consider the concept of

⁶ A position that accords with recent arguments that contend religions represent a form of social niche construction and can consequently be recognised as an extended phenotype (Bulbulia, 2008; Purzycki & Sosis, 2013).

politics as an illustrative example; it is clear that this is not a category of human activity that has always been distinguished as distinct from other spheres of life in all human societies. And it is also true that what would be referred to as politics varies across different eras and in different locations. Yet, this does not make the term semantically incoherent, nor does it make it meaningless to apply to societies that lack a distinct emic concept of politics. There is a clear family resemblance, and a line of continuity, stretching between the recruitment of supporters and formation of alliances in early human hunter-gatherers and the coalition building and campaigning of politicians in modern political systems (Fukuyama, 2011). The argument endorsed in this thesis then is that, just as the concept of politics remains a useful analytic frame the concept of religion is similarly a useful concept when discussing human activity cross-culturally.

The second argument is that although all words carry a specific historical baggage and may have been used problematically in the past, it does not follow that the modern academic usage of a term must mirror such historical usage.⁷ Similarly, a word being etic to a given culture or society does not entail that the concept it describes does not exist or is meaningless to apply.⁸ Although there are certainly many people in the world who are unaware of the concept of demographics, there is no society that does not possess a demographic makeup. An objection to this kind of comparison is that social and cultural

⁷ Responding to criticisms by Josephson (2012) that applying the term religion to discuss Japanese society involved the imposition of an essentialised and politicised category, Ian Reader (2013: 306, 312) noted that such an argument overlooks that modern scholars of religion have acknowledged and moved on from concepts “of an unchanging universal, reified category of religion... founded in nineteenth-century, Euro-American Protestant concepts... and models associated with piety, doctrine, and belief”.

⁸ Consider recognition of the Japanese concept *umami* as a fifth basic taste (Kawamura & Kare, 1987; Lindemann, 2002).

phenomena are intersubjective symbolic constructions and therefore rely fundamentally on the existence of certain shared semantic concepts to be of relevance. The validity of this objection ultimately rests on whether religion is conceived to be a concept like cricket where “you need an explicit set of concepts and norms” for it to exist as a meaningful category, or whether it is a concept like ‘demography’ or ‘economy’, which societies have regardless of whether they have any specific concepts to describe them (Boyer, 2013: 170). The category that religion belongs to depends on whether the term is taken to refer to i) organised religious institutions or ii) a broader and more diffuse array of religious traditions, beliefs, and practices, that do not necessarily involve a specific institutional or hierarchical structure. If the concept is taken to refer to the former, formal institutionalised religion, then it is correct that it is only visible in certain eras and places and it has a distinct monotheistic intellectual history. However, if a broader definition is used it becomes difficult to locate any known society in which some form of religion does not exist. Bulbulia (2004: 661) succinctly summarised this point:

There are few absolutely universal rules in human culture. That religious thought always centres on non-natural or supernatural entities is one of them. The violation of natural expectation is what generates the distinctively sacred quality of supernatural conviction. It is what causes one's neck hairs to stand on end.

Accordingly, the majority of CSR researchers pragmatically adopt something of a neo-Tylorian definition of religion referring “to systems of belief and practice that revolve around commitments to supernatural, anthropomorphic beings (‘gods’)” (Slingerland &

Bulbulia 2011; see also Spiro, 1966b: 96)⁹. Broader formulations also exist; with Slingerland & Bulbulia (2011: 9) describing religion as a “framework of ultimately, non-verifiable metaphysical claims” that does “not necessarily involve[e] explicit theism”. The advantage of such a definition is that it acknowledges that adherence to a specific religious doctrine or institution is not an essential component of religion.¹⁰ The limitation is that it also implies that religion is primarily about metaphysical claims and beliefs and yet for many religious practitioners religion is much more about *practice* than a set of beliefs or theological propositions (McCutcheon, 1997; McGuire, 2008; Pals, 2006; Reader & Tanabe, 1998).

Another definition recently provided by Purzycki et al. (2014: 85) presents a partial resolution to this issue by emphasising that “religion encompasses both religious cognition and behaviour” and noting that these are inextricably linked in a system which “at its most basic level... is composed of ritual behaviours and the supernatural agents to which people appeal”. This definition succinctly addresses criticism of prioritising doctrine over practice by recognising that these are both core elements of religion.

Drawing on the insights and limitations identified with existing definitions, this thesis defines religion as referring to: *the synergistic coupling of ritualised practices and socially ascribed metaphysical beliefs that are orientated towards, or involve interactions with, culturally*

⁹ Such a definition faces the objection that certain iconoclastic and ‘atheistic’ religious systems, like Buddhism, may not fit within the category. However, this argument tends to ignore the everyday reality of religious practices which, even in Buddhist countries, demonstrate clear concerns with supernatural beings (Reader & Tanabe, 1998; Reader, 1990; Samuel, 1993; Spiro, 1982; Tambiah, 1976). The colonial history of how certain traditions became associated with atheism and science (Lopez, 1995, 1998) should also make scholars wary of mistaking an intentionally constructed image with an accurate reflection of a religious tradition’s history and character.

¹⁰ A point that accords with recent findings that the sense of life having an ultimate meaning or purpose appears to be a universal cognitive default regardless of stated personal beliefs (J. L. Barrett, 2012; Bering, 2010; Kelemen & DiYanni, 2005; Kelemen & Rosset, 2009).

postulated superhuman agents and entities. This definition reflects the pragmatic, cross cultural insight of Spiro (1971) and Reader (2004a, 2004b) regarding non-native categories and rejects any a priori association with hierarchical, doctrinal religious institutions. This definition is what should be understood when *religion* is referenced throughout the rest of the thesis, unless otherwise stated.

1.3 Definitions of *Ritual*

The important and diverse roles that rituals play in societies has long been recognised by anthropologists (Douglas 1966; Frazer, 1922 [1890]; Malinowski 1966 [1922]; V. Turner 1969, 1974; Van Gennep 1960; see Kreinath, Snoek, & Stausberg 2007 for an annotated bibliography) and in the CSR field research on rituals is one of the most active areas of research (Boyer & Liénard, 2006; Henrich, 2009; Lawson & McCauley, 1990; McCauley & Lawson, 2002; Schjødt et al., 2013; Sosis & Alcorta, 2003; Whitehouse & Lanman, 2014; Whitehouse, 1995, 2004). Such interest is understandable given that rituals are frequently cited as a human universal (Whitehouse & Lanman, 2014: 675) with Rappaport famously declaring that “no society is devoid of what a reasonable observer would recognize as ritual” (1999: 31).

Yet, despite high levels of interest in ritual there remains little consensus concerning what *ritual* precisely means and whether the term is even meaningful (Bell, 1992; Collins, 2005, 2009; Grimes, 2013; Kreinath, Snoek, & Stausberg, 2006; Whitehouse & Laidlaw, 2004). Part of the issue is the diversity of events and practices that are labelled as rituals; rituals can involve hundreds of thousands of people, like the annual Hindu *Thaipusam* festival in Mauritius (Xygalatas et al., 2013b), but can also refer to individuals performing acts in small

groups or even in solitary, such as the recitation of mantras under isolated waterfalls by Shugendō ascetics (Miyake, 1989). Rituals are typically important elements of religious ceremonies and festivals but they also feature in non-religious contexts, such as sporting tournaments (Boyer & Liénard, 2006; Damisch, Stoberock, & Mussweiler, 2010; Maranise, 2013) and even mundane daily tasks (Hastorf, 2011; Marshall, 2005). Rituals can also be used to mark major rites of passage (Van Gennep, 1960), like the Jewish *Bar Mitzvah* (Zegans & Zegans, 1978), or be performed multiple times a day, such as when Muslims pray facing towards Meccah (*salat*). Rituals can be dramatic and highly arousing, like Shinto firewalking festivals (Lobetti 2013: 37-41), or be austere and muted, such as at weekly Anglicised Christian services. Rituals can last for months, as with the 1000-day meditational ‘marathon’ practices of monks in Mt. Hiei (Rhodes, 1988), or finish in seconds, with the spin of a prayer wheel spun at a Tibetan monastery (D. Martin, 1987; Samuel, 1993). In the face of all of this potential variation, attempting to develop a single definition is daunting and has led some within CSR to assert that the category is too broad to be useful and “we can and should dispense with [it]” (Boyer & Liénard, 2006: 42).

To avoid what they regard as unproductive discussions on the diverse cultural expressions popularly labelled as rituals, Boyer & Liénard (2006) advocate focusing instead on the underlying universal practice of ritualised behaviours. These are described as being “intuitively recognizable by their stereotypy, rigidity, repetition, and apparent lack of rational motivation” and are noted to be active “in a variety of conditions, customs, and everyday practices... [including] children’s complicated routines [and] in the pathology of obsessive-compulsive disorders (OCD)” (Boyer & Liénard 2006: 1). Boyer & Liénard argue

that what links these diverse areas is the activation of cognitive ‘Hazard-Precaution’ and ‘Action Parsing’ systems, which serve to focus human attention towards “the detection of and reaction to inferred threats to fitness” from the environment, especially sources of contamination and pollution (2006:1).

Nevertheless, despite Boyer & Liénard’s protests against the usefulness of the category of ritual, the very article in which they outline this position also serves as an illustration of why the term is necessary. Specifically, Boyer & Liénard repeatedly switch between using the terms ritual and ritualized behaviours despite maintaining that the two are conceptually distinct categories. Hence, the title of the sub-section containing a list of “general features of ritual” is within the section described as “different features of ritualized behavior” (2006: 4)¹¹. From their response to commentaries on the article is clear that they intended the list to refer to ritualized behaviours but regardless of intent the general features they provide are equally applicable to ritual more generally.

The features they outline are: compulsion- “people feel that it would be dangerous or unsafe or improper not to perform ritualized actions”; rigidity/adherence to script- “people feel they should perform a ritual in the precise way it was performed before”; goal demotion- “performance is divorced from observable goals”; internal repetition and redundancy “of the same action or gesture, as well as reiterations of the same utterances”; and a restricted range of themes, commonly relating to “pollution and purification, danger and protection” (Boyer & Liénard, 2006: 4).

¹¹ Several invited commentaries on their article make similar observation, noting that cultural rituals are conflated with ritualised behaviour when engaging in broader discussions of their implications of their theory (McCauley, 2006; Merker, 2006; R. Turner, 2006; Wang & Bello, 2006).

These characteristics are drawn from a wide range of sources and individually each does appear important to the general concept of ritual, however while the first four characteristics enlarge the reference category and enable the inclusion of non-religious events, conversely the fifth excludes a substantial portion of the ritual practices visible around the world. In particular, celebratory rituals that lack significant reference to themes of purity and pollution but are prolific the world over are excluded (Merker, 2006). Furthermore, competitive ritual events which involve performances that are coherently linked to desired goals (Abbink, 1999) problematize the goal demotion characteristic. Rather than understanding these characteristics as required features for something to be recognised as *rituals*, it would thus seem more effective to interpret them as being core features of a family resemblance. Each characteristic is indicative of ritual but none is ultimately required for a practice to be categorised as a ritual.

Boyer & Liénard are keen to disentangle their interest with ritualized behaviour from any particular association with religion but conversely Roy Rappaport, one of the most influential scholar of rituals¹², argues that ritual represents the very “ground from which religious conceptions spring” (1999: 3). However, despite noting this connection, Rappaport favours a more inclusive definition for ritual describing it as “the performances of more less invariant sequences of formal acts and utterances not entirely encoded by the performers” (1999: 3, 24). The key features in this account are that rituals represent ‘conservative patterns of formalised actions’ and that the sequences of actions involved are not

¹² Boyer & Liénard reference Rappaport as the primary source for the five features of ritual.

understood as having been created, or authored, by any individual ritual performer. Neither of which introduce any irreducible *religious* elements.

The influential cognitive anthropologist Harvey Whitehouse has expressed some scepticism about the utility of broad conceptual categories, however despite this he has also outlined two recurrent features found in most phenomena described as ritual: 1) a historical/mythical origin and 2) casual opacity (Whitehouse 2004: 3-4). Incorporating aspects of Rappaport's definition and work by Humphrey & Laidlaw (1994: 5), Whitehouse notes that ritual participants "both are, and are not, the author of the acts" as ritual actions performed are accorded a past that places them "outside the mind of the actor", linking them "to a procession of shadowy predecessors, primordial ancestors, or gods" (Whitehouse 2004: 4). This emphasis means that rituals are almost always presented as ancient, even when they are recent inventions. Second, rituals invariably involve casually opaque procedures, which cannot be reduced to instrumental factors (Whitehouse & Legare, 2012). Causally opacity is typically defined as relating to actions which are not linked by a coherent chain of "physical-causal mechanisms or processes" (Legare et al. 2016: 86). For instance, in Europe there is a well-established folklore concerning 'wishing wells' that will grant wishes when offered coins. Yet when throwing a penny into a well there is no intuitive physical-mechanical chain, nor a culturally mandated mechanism, to explain how this will enable the wish to be granted. Wishing wells thus involve a *casually opaque* ritual procedure despite also possessing a coherent instrumental goal (the granting of a wish). Causal opacity thus does not entail that rituals lack an instrumental purpose it is rather that they are not expected to achieve that purpose "by means that are transparent or

even in principle knowable based on principles of physical causality” (Legare et al. 2016: 87).

There is however occasionally confusion on the distinction between levels of opacity in the literature as illustrated by Kapitány & Nielsen (2015: 15) recently describing causal opacity as being “when we cannot easily infer the *instrumental purpose*, causal mechanism, or even the performer’s *underlying motivation*” and misleadingly indicated that it “is defined as a situation in which there is start-end state equivalency”. Start-end equivalency is a scenario which can generate causal opacity (Legare, Wen, Herrmann, & Whitehouse, 2015) but it is not a prerequisite for opacity to exist. But more problematic is the argument that causal opacity relates both to higher level goals (a ritual’s purpose or the performer’s motivation) and lower level causal sequences. Elsewhere in the literature various scholars, including Kapitány & Nielsen (2016), have affirmed the importance of this distinction by using ‘goal demotion’ to refer to the ability to infer and understand the reason (goals or motivations) for a given action sequence (Schjoedt et al. 2013: 45). This distinction is important as goal demotion and causal opacity do not always occur simultaneously in rituals, as illustrated in the wishing well example.

Consequently, although the first feature of uncertain authorship identified by Whitehouse is a core aspect of almost all ritual performances, the prevalence of the second feature of casual opacity depends on whether it refers to action sequences lacking intuitive physical-causal mechanisms or the broader definition incorporating goal demotion. Under the more restrictive definition it can be seen as a universal feature of ritual but the second fails to offer an account for rituals which invoke supernatural causality or involve coherent

instrumental outcomes, such as ritualized status competitions involving structured combat with stylised weapons (Abbink, 1999).

For instance, although the mechanism by which ritual offerings of food provided to various supernatural beings in Southeast Asia and elsewhere is ‘consumed’ by such entities is unknown, the instrumental *purpose* of providing food is often clearly understood. The goal of such actions are understood to be about gaining blessings, avoiding harm, or curing an illness. It is also widely recognised that the supernatural beings being propitiated can be the source of misfortune and fortune. Whether it is accurate to describe ritualised offerings as being causally opaque then depends on whether we focus on lower or higher levels of analysis and whether we acknowledge causal pathways that invoke supernatural agency as offering logical causal connections (Schjødt et al., 2013; Sørensen, Liénard, & Feeny, 2006; Sørensen, 2007). Ultimately, as with the characteristics identified by Boyer & Liénard, the most profitable approach would seem to be to regard the characteristics identified by Whitehouse as being indicators of a family resemblance for rituals.

The final approach to defining rituals to be examined is that presented recently by Rossano (2012) which outlines a detailed typology of ritual. Rossano supports Boyer & Liénard’s call to distinguish between ritualised behaviour and cultural rituals but does not agree that purity and pollution are fundamental to this distinction. Referencing Tinbergen (1952), he notes that *ritualization* has long been defined by ethologists to refer to “an evolutionary process whereby an incidentally informative behaviour becomes isolated and specialized to be socially informative”, and that ritualized behaviour is “widespread in the animal world and essential for regulating social interactions” (2012: 532). For instance,

ritualized attack displays, submissive postures, and mating ‘dances’ all illustrate the capability of animals to emancipate certain informative elements from their associated full behavioural expression (Armstrong & Wilcox, 2007).¹³ In humans, a comparative example of early developing ritualised behaviour is when young children extend and raise their arms, in mimicry of a climbing movement, in order to elicit a caregiver into picking them up (Rossano, 2012: 532). This shortening process has been referred to as *ontogenic ritualization* and does not require complex cognitive abilities to perform (Tomasello, 2008: 22-23).

Reflecting many of the themes already discussed, Rossano suggests that ritualized behaviour is defined by three core characteristics: i) the emancipation and formalization of some reduced element from a longer behavioural sequence; ii) rule governance, in that ritualised behaviours must be performed in a prescribed manner, and iii) repetition, the behaviours are often repeated to attract and hold attention (Rossano, 2012: 532). An additional fourth characteristic of goal demotion is also identified as being relevant to human specific ritualized behaviours, mirroring the emphasis on causal opacity discussed by Whitehouse & Legare (2012).

In contrast to ritualised behaviour, ritual is defined by Rossano as “referring to a variety of scripted, ceremonial and symbolic activities” that are *only* found amongst humans. Rituals “include ritualized behaviors, but [also] other important features... such as sacredness, symbolism, traditionalism... performance” and a magnification of scale (2012:

¹³ Many animals, instead of engaging in direct, potentially deadly, combat will perform ritualized attack displays, in which patterns of behaviour associated with attacks are performed without any accompanying attack taking place. Such displays are often sufficient to elicit a submissive response from the target (Givon, 1995).

532). Symbolism here refers to the ability to attach distinctive meanings to objects and behaviours that allow them to represent something else, such as a simple pattern on a flag serving as a symbol of a nation. While there is evidence of some limited symbolic capabilities in a few other species, humans display a remarkable symbolic capacity from a young age. Indeed, the bridge between ritualised behaviours and broader cultural rituals is to be found in the interactions between children and caregivers (Rossano 2012: 533-539)¹⁴.

The developmental research illustrates the importance of ritualized behaviour to human development and demonstrates how it becomes increasingly elaborated as children mature. Short ritualized interactions and simple games are replaced over time with verbal and demonstrative introductions and these in turn give way to elaborated family and community cultural ritual practices. Rossano argues that it is the elaborative and norm inflected character of these later cultural rituals that generate the unique aspects of human rituals. Forging strong emotional bonds that can incorporate non-kin, and serving as vehicles to transmit and reinforce group values and norms.

Rossano's broader arguments about the function of rituals are not yet validated but the typology of ritual that he outlines is coherent, empirically supported, and avoids the limitation of themes imposed by Boyer & Liénard. This thesis agrees with the divisions

¹⁴ Early infant-caregiver interactions almost immediately display all four of the criteria of ritualized behaviour. Young children display an ability to quickly acquire the normative structure of games (Rakoczy & Tomasello, 2009; Rakoczy, Warneken, & Tomasello, 2008, 2009) and a tendency to protest violations of perceived norms or rules (Mesman, van IJzendoorn, & Bakermans-Kranenburg, 2009; L. Murray & Trevarthen, 1986; Nadel, Carchon, Kervella, Marcelli, & Réserbat-Plantey, 1999; Stormark & Braarud, 2004; Tronick, 2003), even in the absence of explicit instruction (Schmidt, Rakoczy, & Tomasello, 2011). They also willingly imitate causally irrelevant actions when they recognise intentionality behind an adult's actions (Brugger, Lariviere, Mumme, & Bushnell, 2007; Gergely, Bekkering, & Király, 2002; Gergely & Csibra, 2005), and ritualized behaviour appears to serve as a signal that stimulates over-imitation (Lyons, Damrosch, Lin, Macris, & Keil, 2011; Lyons, Young, & Keil, 2007; Nielsen et al., 2014; Nielsen & Tomaselli, 2010; Whitehouse & Legare, 2012).

proposed and seeks to adhere to Rossano's typology. However, when referring to *rituals* without qualification, the term should be understood not to reference any single prototypical example but rather to be based on a family resemblance between practices that are i) *rule or norm governed*, ii) include *repetitive behaviours*, iii) involve actions that are *causally opaque*, iv) invoke mythological, ancestral, or *uncertain authorship*, and v) contain *substantial symbolism*. No individual feature is essential for something to be considered a ritual, but the more that are present the stronger the case for inclusion in the category. Rituals under such a definition do not necessarily invoke supernatural beings or relate to a restrictive range of themes.

The arguments against the utility of retaining the term, as with the arguments against *religion*, are also rejected on the same grounds- that any useful analytical term inevitably requires both abstraction and the imposition of some form of artificial separation. No value judgment is intended by referring to a practice as a ritual, since it is evident that all societies possess a large repertoire of cultural rituals.

1.4 Ritual as a 'Costly Signal'

The performance of ritual acts that are costly in terms of their physical or emotional toll and the amount of time and effort involved presents a psychological and evolutionary puzzle. On a psychological level, the key question is what processes are motivating people to participate, whereas on a broader evolutionary level the question is how individuals and groups practicing such costly behaviour are able to survive.

Costly signalling theories build on insights from evolutionary biology that traits can be selected for that result in a 'handicap' (Zahavi & Zahavi, 1997; Zahavi, 1975) to an

individual when they serve as a reliable signal of fitness (e.g. a peacock's tale). Applying this logic to religion and ritual, researchers (Irons, 2001; Sosis & Bressler, 2003) have suggested that participation in cultural rituals, particularly those that are traumatic or painful, present a hard to fake signal of commitment to an ingroup and simultaneously demonstrate an individual's quality as a potential partner. Participating in dangerous, painful, or expensive rituals thus serve as public signals that individuals are willing and have the physical or emotional resources to endure the cost of ritual participation.

Psychological efforts to address why individuals participate in and hold positive regard for groups with harsh practices have primarily centred on the concept of *cognitive dissonance*. Cognitive dissonance theory, as introduced by Leon Festinger (1962), suggests that people have a compulsion to maintain harmony in their beliefs, attitudes, and actions. This was famously demonstrated in a study by Aronson and Mills (1959) that revealed individuals reported a stronger liking for groups that involved more severe initiations. Although subsequent research has overall confirmed the importance of dissonance reduction in determining individuals' responses to groups with harsh initiations (Gerard & Mathewson 1966; Hautaluoma & Spungin 1974; Kamau 2013), the theory does not help to resolve questions such as what motivates individuals to choose to join groups with harsh initiations, why such groups persist, or the psychological effects on those who are part of the group but do not participate in such performances.

Costly signalling theories can help to address these gaps. For example, Henrich (2009) argued that costly practices, including costly rituals, are an important subset of a more general category of credibility enhancing displays (CREDs) which help learners to

determine “how much to commit to a particular culturally acquired mental representation” (Henrich 2009: 245). CREDs are defined as actions that are consistent with verbal representations expressed by an individual, particularly actions that “would not be performed by a model believing something different from what [was] expressed symbolically” (Henrich 2009: 247). Henrich elaborated that a sensitivity to CREDs is a crucial component of the processes that underpin social learning as it enables learners to avoid exploitation from insincere manipulators. Thus, Henrich inserts costly rituals into an evolutionary framework in which intergroup competition and cultural group selection processes selected for groups with interlocking systems of beliefs and costly practices. Specifically, groups with costly practices “that elevate the commitment of group members to beliefs that promote group benefits, large-scale cooperation and solidarity, and... success in competition with other social groups” were favoured (2009: 245).

Henrich’s account offers a broader evolutionary logic as to why counterintuitive religious beliefs are almost always paired with costly ritual practices, and locates the origins of a fascination with costly displays in core processes of social learning. He contends that this distinguishes his account from earlier treatments of costly rituals, which he summarises as lacking an elaborated evolutionary model and positing that committed believers “will be able to perform the costly requirements more cheaply than nonbelievers” (Henrich 2009: 256).

Henrich is correct that his emphasis on social learning processes distinguishes his approach from other signalling theories but his wider claims that earlier models lacked an evolutionary framework and were focused on costs being ‘cheaper’ for believers are less

well supported. For instance, even in Irons' pioneering work (Irons, 1996a, 1996b, 1996c, 2001) exploring ritual and religion from a costly signalling perspective, there is a clear evolutionary framework and reference to cultural group selection processes. Hence, Irons argued that religion, costly rituals, and taboos all function as credible 'costly to fake' commitment signals (Zahavi & Zahavi, 1997) that enable greater intragroup cooperation. Religious beliefs that signal a particular group identity but involve only low-cost behaviours could easily be faked but when a religion contains teachings that require adhering to restrictions and participation in costly ritual practices, faking commitment becomes more difficult. This accords almost precisely with the account presented in the CREDs model (Henrich, 2009; Lanman, 2012; Wildman & Sosis, 2011).

Sosis and Bressler's (2003: 232) costly signalling account, which Henrich specifically highlights as offering an alternative to his model, does argue for a distinctive quality to religious rituals that are orientated towards sacred "nonmaterial beliefs that cannot be falsified" that makes them the "most effectual at elevating group cohesiveness". However, their model draws heavily on Irons' work and hence although they regard religious rituals as uniquely potent they also explicitly acknowledge a broader adaptive evolutionary framework in which costly signals, including rituals, function as a means of enhancing intragroup cooperation and cohesion and increasing intergroup fitness.

Consequently, it seems that the CRED account has more in common with earlier theories than Henrich suggests. More recent formulations have incorporated the insights of Henrich and earlier theorists and have argued accordingly that costly rituals represent a "social technology" (Fischer & Xygalatas, 2014) for establishing higher levels of cooperation

within a group (Soler, 2012). If these interpretations are correct, then this 'technology' could provide extreme ritual groups with a competitive advantage in intergroup scenarios (Henrich, 2009; Wildman & Sosis, 2011).

Support for costly signalling theory in a ritual context is primarily drawn from theoretical treatments (Bulbulia & Sosis, 2011) and modelling studies (Henrich, 2009; Wildman & Sosis, 2011) but there are some limited studies based on real world data. Sosis & Alcorta (2003), for example, analysed records of 19th Century religious settlements in the US, and found that religious communes with the most costly ritual requirements survived longer than both secular communes and other religious communes with less costly requirements. Similarly, examining Indonesian official statistics collected during the recent financial crisis, Chen (2010) found that households which intensified their religious investments experienced an almost 50% reduction in their need for welfare or loans four months later. In comparison, households that *reduced* their religious activities, or did not participate at all, saw smaller decreases in loans of 21% and 6% respectively. Finally, a recent field study, conducted during a Hindu festival in Mauritius, found that after observing an extreme ritual, there was a trend for observers to donate more to their temple than the performers, although this difference did not reach significance (Xygalatas et al., 2013b). What all of these examples share in common is that the costs incurred by ritual participation were translated to benefits at both the individual and group level.

These cases seem to offer persuasive illustrations but it is important to recognise that in all cases there is always the possibility of unaccounted for confounding variables. Moreover, they also deal with evidence pertaining to different levels of analysis; ranging

from the individual and the household to entire religious communities. Thus, it would be wrong to consider all of the studies as offering support for a single unified theory of costly signalling. Rather, it is more appropriate to consider costly signalling as representing a general evolutionarily-informed approach that contains a constellation of theories that offer a diverse range of predictions that are united by a concern with actions and beliefs acting as 'signals' to potential allies or sexual partners.

An illustrative example of why it is important to consider the diversity of signalling theories is provided in a large scale coding analysis conducted by Sosis, Kres, & Boster (2007). This study coded the 60 societies included in Probability Sample File (PSF) from the Human Relations Area Files (HRAF) and explicitly sought to distinguish between competing signalling hypotheses. Their results demonstrated evidence that "males in societies that engage in warfare endure the costliest rites" and that whether a society fought in wars against other cultural groups was "an important determinant of whether or not male rites result in permanent visible marks, such as ritual scars" (Sosis et al., 2007: 234). This provided support for signalling theories that focus on intergroup conflict and commitment but they found no support for predictions that costly male rites might serve as a signal of mate quality or correlate with the intensity of mating competition.

In addition to issues arising from the diversity of signalling theories, some may object that such approaches are too explicitly concerned with social interactions to qualify as a cognitive approach to ritual. But this position is contradicted by modern cognitive research which recognises that many of our core cognitive processes are orientated to both emotional reactions (Damasio, 2008; Haidt, 2001) and social connections (Robin I. M.

Dunbar, 2003; Lieberman, 2013). As such, addressing the effects of ritual performances on both performers and observers should be a core issue for any comprehensive cognitive theory of ritual.

In summary, costly signalling theories do provide potential answers to the questions omitted by dissonance accounts by arguing that costly rituals can serve as indicators of a group's quality and the commitment of its members and thus make a group more attractive to join. They also suggest that increased positive regard for groups with costly rituals can be translated into behavioural support and community cohesion and increase the longevity of the group. However, less is known about how observing rather than participating in costly rituals affects group members but there is some preliminary evidence that it increases generosity (Xygalatas et al., 2013b). Ultimately, costly signalling theories' greatest strength is their ability to generate testable predictions and they thus serve as an important reference point in the empirical chapters that follow.

1.5 The Divergent Modes of Religiosity Theory

The next cognitive theory of ritual to be examined in detail is the *Modes of Religiosity*¹⁵ developed by the cognitive anthropologist Harvey Whitehouse (Whitehouse & Lanman, 2014; Whitehouse, 1992, 1995, 1996, 2002b, 2004). This theory has been extremely influential from the inception of the CSR field, has inspired a large body of theoretical and empirical research, and specifically addresses high arousal rituals, all of which makes it of crucial significance for this thesis. In response to new findings, the theory has also undergone

¹⁵ Referred to from this point on as 'Modes theory' or simply 'Modes'.

significant revisions in recent years and it is thus crucial to clarify what predictions are made in the most recent formulations.

Modes theory offers a selectionist account (Sperber, 1994) of why certain patterns of religious behaviour appear universally recurrent, identifying a variety of cognitive mechanisms that generate enduring memories and instil motivation to transmit beliefs and knowledge to others (Whitehouse, 2004: 63-83). The theory primarily addresses a “divergence in modalities of religious experience and practice” (Whitehouse, 2004: 63), between *imagistic* small scale local traditions, with loosely defined doctrines and highly arousing but rarely performed rituals and larger, more hierarchical *doctrinal* traditions that contain explicit doctrines, transmitted through frequently repeated low arousal rituals. This distinction represents a phenomenon long recognised in research on religion, and featured in the works of many pioneering social scientists including Weber (1930) and Durkheim (1965). In anthropology typologies that recognised similar variations in religious groups and their practices were presented by influential scholars including Ruth Benedict (1934), Victor Turner (1969), Melford Spiro (1982), Iona Lewis (1989), and Roy Rappaport (1999). These theories remain influential in anthropology but due to the ‘reflexive turn’ (Clifford & Marcus, 1986; Rabinow, 1977; Spiro, 1996) theories that are broader and cross cultural in scope are now often criticised for being overreaching, failing to engage properly with *embodied* or *lived* practices (Csordas, 1990; Lock, 1993), and ignoring the distorted image that is generated by the influence and privilege of the anthropologist (Grimes, 2013; Kreinath et al., 2006). Although Whitehouse did not cite these arguments, he still seeks to distinguish the Modes theory (2000,2004) from the weaknesses of these earlier anthropological accounts

by combining detailed ethnographic knowledge with insights about human cognition derived from modern cognitive science.

Modes theory originated in Whitehouse's ethnographic work (1995) conducted with the Baining people of Papua New Guinea in the 1990s. Initially, he set out to explore the traditional hunting methods and gardening techniques practiced in the region but quickly shifted his focus after he found that people were much more interested in discussing their religious beliefs- "something much closer to their hearts" (Whitehouse, 2015). The majority of Baining belonged to the *Pomio Kivung* a millennial cargo cult that emerged in the 1950-60s. By the 1990s this group had developed into a large mainstream organization, with a well-defined hierarchy promoting a specific orthodoxy of views. However, during Whitehouse's fieldwork a small-scale, splinter movement emerged that encouraged innovative interpretations of the core Pomio Kivung religious ideology and employed rituals that were dramatic and emotional. These two modes of practice stood in sharp contrast to one another but Whitehouse recognised them to be divergent manifestations of *the same underlying religious system*. He also argued that this was not an isolated development but was both common across Papua New Guinea and he maintained represented a recurrent global pattern (Whitehouse, 1995, 2000).

This broader prediction was initially based on the hypothesis that the divergent modes of religiosity (Table 1.1) were created through the activation of distinctive systems within explicit memory. Specifically, Whitehouse proposed that the dramatic, often terrifying, high arousal rituals of the imagistic mode generated rich, episodic *flashbulb* memories (Conway, 1995). These kind of memories encode experience as specific narrative events,

demonstrated by the way that people remember transformative events, like weddings or deaths. Alternatively, for the low arousal, repetitive rituals of the doctrinal mode, which lacked the emotional impact of the imagistic rituals, non-specific semantic and implicit behavioural memories were generated. The central argument presented was that in addition to the contrasts in emotional intensity there was also a clear distinction between both modes in terms of the *frequency* of ritual performances (Fig.1.1), and this had significant implications for the memorisation and transmission of religious systems.

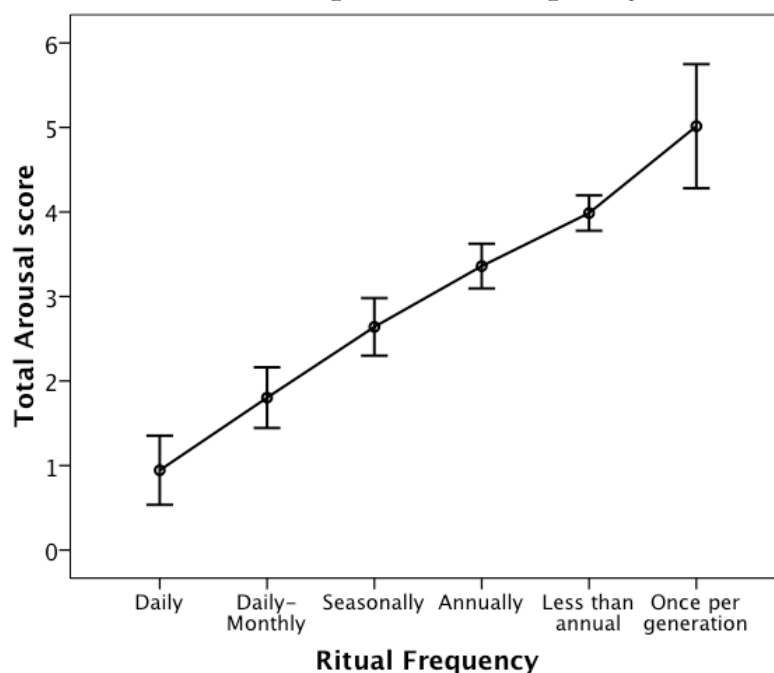
Table 1-1: Features of doctrinal & imagistic ‘modes’ of religiosity

	Doctrinal	Imagistic
Psychological Features		
1. Transmissive Frequency	High, frequently performed	Low, rarely performed
2. Level of Arousal	Low	High
3. Principal Memory System	Semantic Schemas and Implicit Scripts	Episodic/ Flashbulb Memory
4. Ritual Meaning	Learned/Acquired, Doctrinally sanctioned and prescribed	Internally Generated through spontaneous reflection (SER)
5. Techniques of Revelation	Recognised leaders and authoritative texts employing rhetoric, logical integration and narratives.	Collective participation in rituals results in a multi-vocal and multivalent interpretation of icons and symbols.
Socio-political Features		
6. Social Cohesion	Diffuse across large ‘imagined communities’	Intense within localized communities
7. Leadership	Highly elaborated hierarchy with ranked levels, typically centralised and institutionalised	Largely passive, but where present charismatic leaders exercise authority personally
8. Inclusivity/Exclusivity	Inclusive	Exclusive
9. Spread	Rapid, Efficient	Slow, Inefficient
10. Scale	Large Scale	Small Scale, Localized
11. Degree of Uniformity	High; largely homogenous	Low; possibly fragmented among multiple traditions
12. Structure	Centralized and institutionalised	Loosely organised and diffuse, poorly defined

Features adapted from Atkinson & Whitehouse, 2010; Martin, 2004: 11; Whitehouse, 2004: 74.

To elaborate, in the doctrinal mode, ritual performances tend to occur frequently but are conducted in a highly repetitive or routinized format, often accompanied by a calm and austere environment or atmosphere. Doctrinal rituals also tend to include some verbal transmission of complex theology and doctrine. The repetitive nature of this combination of low arousal ritual and doctrinal exposition generates strong implicit memories and ALSO aids in the formation of semantic memories for both ritual procedures and complex doctrinal exegesis (Whitehouse, 2004: Ch 5). Illustrations of the doctrinal mode abound within the major world religions and they tend to be associated with centralised, ranked hierarchies, professional priesthoods, and large ‘imagined’ communities of believers (B. Anderson, 1983).

Figure 1-1: Arousal level compared with frequency in 645 rituals



Source: (Atkinson & Whitehouse, 2011: 55)

One issue that arises from the frequent repetition of doctrinal rituals, along with the presence of a developed official doctrine, is that this package tends to provide little

motivation for independent reflection and risks producing a *tedium effect*: “a state of low morale arising from overfamiliarity with religious formulae and routine” (Whitehouse, 2004: 98). Whitehouse suggests that as a consequence of this, the doctrinal mode possesses an inherent structural paradox that, in the right circumstances, can encourage the emergence of dramatic imagistic splinter groups. Even though it is possible for such splinter groups to establish an independent tradition, the evidence from Melanesia suggests that it is more common for them to be reabsorbed and serve to reinvigorate the main doctrinal tradition (Whitehouse, 2004: 133; but see McCauley & Lawson, 2002: 202). In this sense, Whitehouse presented the disjunction between imagistic and doctrinal modes as being potentially complementary and strengthening rather than a non-porous dichotomy. Yet, he also cautioned that traditions which try to maintain both forms of ritual practice are likely to represent a less stable arrangement.

In contrast to the sedate rituals of the doctrinal mode, the imagistic mode is characterised by intense, rarely performed rituals, which often invoke high levels arousal via “extreme forms of deprivation, bodily mutilation and flagellation, and psychological trauma based around participation in shocking acts” (Atkinson & Whitehouse, 2010: 51). Imagistic rituals are paradigmatically illustrated by Melanesian ‘rites of terror’ (Whitehouse, 1996) but can also relate to more euphoric experiences such as the ecstatic trances of Amazonian shamans (Vitebsky, 2001). These kinds of dramatic experiences tend to generate enduring memories, which Atkinson & Whitehouse attribute partly to the “considerable evidence linking pain and trauma to the activation of durable episodic recall” (2011: 57). Imagistic rituals rarely involve any explicit doctrinal transmission, instead it is

collective participation and the subsequent independent reflection, termed *spontaneous exegetical reflection* (SER) in Modes theory, which generates religious knowledge and distinguishes members as belonging to particular communities. Independent reflection in such circumstances is motivated by a lack of any officially prescribed sense of *meaning* to the rich, affective, and personally consequential episodic memories. Due to ritual participation defining the boundaries of imagistic communities and the intense relational bonds formed between ritual performers, this mode of religion tends to be harder to spread and is likely to remain localized.

The significance of these divergent modes of ritual have been examined across a number of articles and edited volumes that have tested the predictions of Modes theory against a range of ethnographic (Whitehouse & Laidlaw, 2004), historical (Whitehouse & Martin, 2004), and psychological data (Whitehouse & McCauley, 2005). Although generally the studies to date have affirmed the importance of the distinction Modes theory posits, a persistent criticism is that in practice the two modes are rarely entirely distinct. Several researchers have commented that all of the worldwide doctrinal religions contain local variations which fall closer to the imagistic spectrum and that there are also various syncretic traditions which incorporate aspects of both modes (Whitehouse & Laidlaw, 2004; Whitehouse & Martin, 2004; Wiebe, 2004). Stausberg, for instance, when responding to a recent formulation of Modes theory (Whitehouse & Lanman, 2014), contended that “rather than dichotomizing ritual into two categories, a fuller ecological scenario of societies would probably need to take account of a broader spectrum of ritual layers, modes, performances, or packages” (Stausberg, 2014: 689).

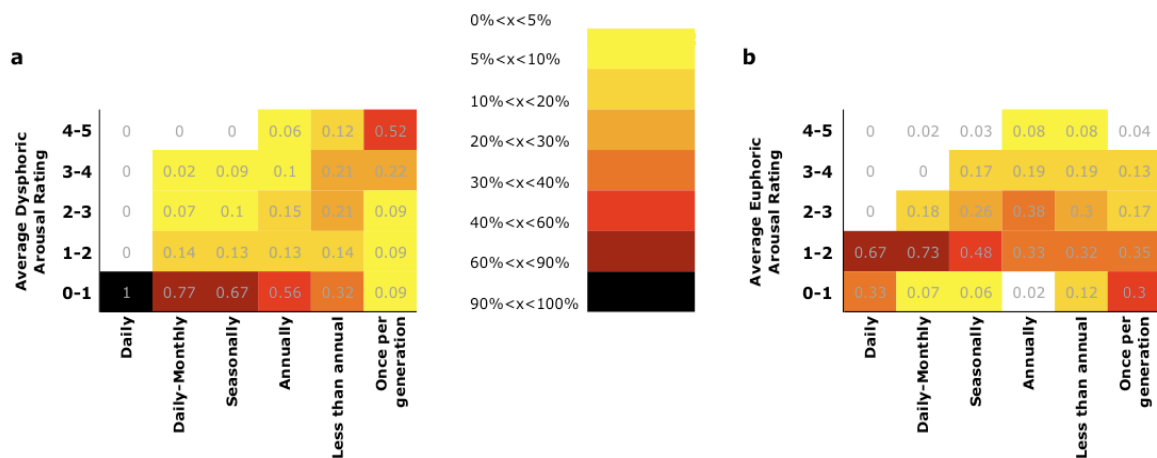
Whitehouse has responded to these criticisms by arguing that he did not reject the possibility of imagistic and doctrinal practices co-existing in certain scenarios¹⁶. He also acknowledged that such cases are not rare but contends that this is due to “the doctrinal mode [being] inherently vulnerable to the tedium effect”. Accordingly, in addition to occasionally generating imagistic splinter groups, Whitehouse recognises that “more sensually evocative forms of religious activity” can also develop within a doctrinal tradition (Whitehouse & Martin, 2004: 217). However, he argues that in such cases there remains an important distinction in how epiphanic revelations from high arousal practices are responded to. In particular, in doctrinal conditions they “tend to be subject to frequent, subsequent rehearsal as more or less formulaic narratives” and this inevitably reduces creative interpretations and lessens their potential innovative impact on the tradition (Whitehouse & Martin, 2004: 217).

Overall, despite persistent accusations of enforcing a rigid binary between the modes, Whitehouse has repeatedly admitted that “there is no such thing as a religious tradition which does not depend for its transmission on both semantic and episodic memory” (Whitehouse, 2000: 11) and that the divergent modes should thus not be understood as representing a dichotomous typology but rather as two *cognitively attractive* poles of a ritual spectrum (Atkinson & Whitehouse, 2011). The prediction of the Modes theory then is not that there will be no groups that have ritual practices combining imagistic and doctrinal elements but rather that such groups are unlikely to be as stable or successful as groups

¹⁶ Indeed, as discussed previously, his original ethnography presented an account in which an imagistic splinter group was eventually reabsorbed into a mainstream doctrinal tradition.

that adhere *predominately* to one of the two modes. Despite the fact that there is now a collection of edited volumes addressing Modes theory (McCauley & Whitehouse, 2005; Whitehouse & Laidlaw, 2004, 2007; Whitehouse & Martin, 2004) and many cases have been offered in support of the theory, a clear limitation of such evidence is that it could be due to cherry picking cases that accord with the theory.

Figure 1-2: Heat map of (a) dysphoric and (b) euphoric arousal against frequency



Note: Darker colours represent more cases

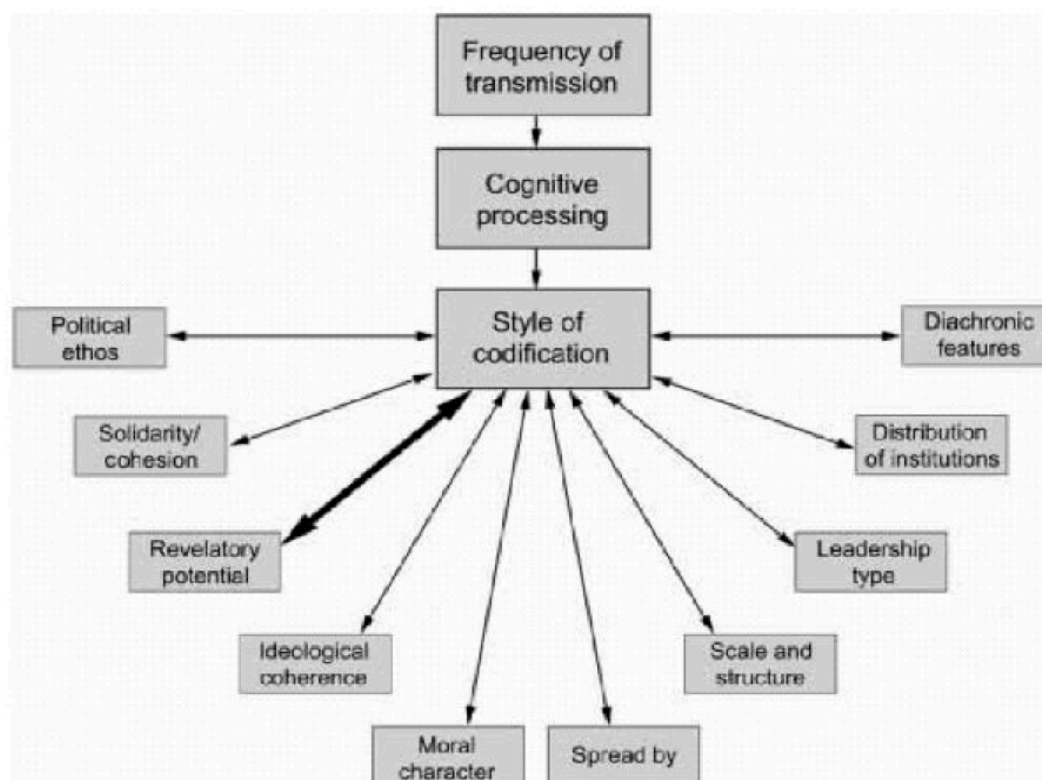
Source: (Atkinson & Whitehouse 2011: 56)

To help resolve this issue empirically a large scale coding analysis of ethnographic material from the HRAF database was conducted that examined 645 rituals from 74 cultures worldwide (Atkinson & Whitehouse, 2011). This study confirmed the existence of a clustering of ritual forms around the high frequency, low arousal and low frequency, high arousal poles (Fig.1.2), and also that the “emotional arousal experienced by ritual participants [was] inversely correlated with ritual frequency, such that less frequent rituals tend to elicit higher emotional arousal ratings” (Atkinson & Whitehouse, 2011: 56) (Fig.1.1). Thus, the core predicted patterns of ritual performance and frequency were found to exist

from a more representative sample of rituals. It is important to note here that this analysis addressed specific rituals and not religions, thus the data demonstrated only that rituals clustered in the predicted manner, not religious traditions.

Another important criticism of Modes theory, levelled most prominently by McCauley & Lawson (McCauley & Lawson, 2002; McCauley, 2001), is that the theory amounts to attributing all features of a religion to the demands of long term memory and results in differential ritual frequency serving as the single causal variable for all aspects of a religious tradition (Fig.1.3).

Figure 1-3: Direction of influence among the variables of Modes theory



Source: (McCauley & Lawson, 2002: 105)

This criticism has greater validity in relation to earlier formulations of the theory (Whitehouse, 1992, 1996), however more recent versions are less guilty of providing a

mono-causal explanation. Instead, the two most recent book length elaborations of Modes theory (Whitehouse, 2000, 2004) detail a complex constellation of factors that influence the political scale, structure, and ethos of religious communities, including surrounding socio-political factors (Table.1.1) and cognitive biases. The effects of the various factors also have been typically represented as interactive as opposed to unidirectional (Whitehouse, 2004). As a result, more recent formulations of the theory have sought to examine how surrounding societal and environmental factors, such as the prevalence of intergroup conflict and complexity, can impact the dominant ritual form (Atkinson & Whitehouse, 2011; Whitehouse & Hodder, 2010) and how ritual mode relates to distinctive affiliative processes (Swann et al., 2012; Whitehouse & Lanman, 2014; Whitehouse, 2015). Consequently, criticisms that Modes theory is monocausal (McCauley & Lawson, 2002) and rigidly dichotomous (Grimes, 2013: 200; Stausberg, 2014) appear to be overstated, at least in reference to more recent formulations of the theory. Yet, despite this the differential frequency of performance and associated mnemonic processes do remain crucial components of the theory and should not be overlooked.

In addressing imagistic rituals Modes theory draws on research relating to the formation of vivid flashbulb memories (FBMs), which tend to be generated by events that are personally consequential and traumatic (R. Brown & Kulik, 1977; Conway, 1995). Yet, flashbulb memories are still a contentious topic in memory studies, with long enduring debates over whether they rely on *special* or ordinary mnemonic processes (Bohannon, 1988; Christianson, 1989; Davidson & Glisky, 2002; Lanciano & Curci, 2012; Luminet & Curci, 2008). There are also disagreements about whether they are more accurate and enduring

than ordinary memories or simply engender greater confidence (Conway, 1995; Cubelli & Della Sala, 2008; Talarico & Rubin, 2003, 2007; Winograd & Neisser, 2006). A recent ten-year study of memories of the September 11th attacks, for instance, found that “there was rapid forgetting of both flashbulb memories within the first year... [and] inconsistent flashbulb memories were more likely to be repeated rather than corrected over the 10-year period” (Hirst et al. 2015: 604). The authors of this study then concluded that “even traumatic memories and those implicated in a community’s collective identity may be inconsistent over time and these inconsistencies can persist” (Hirst et al., 2015).

Whitehouse acknowledged the debates concerning flashbulb memories in an earlier review of the literature (2004: 105-117) but concurred with Conway’s (1995) view that the critical literature had “more or less ignored Brown and Kulik’s original claims about the nature of [flashbulb memories] and has consequently failed to test them” (Whitehouse, 2004: 108). Specifically, critical studies have tended to focus on demonstrating the “striking inaccuracies in FBM recollections”, highlighting that since they “are subject to distortions and forgetting... they cannot be regarded as special memories” and occasionally concluding that “FBMs do not exist” (Curci & Conway, 2013: 352). Whitehouse (2004: 105-117) argues that this approach overlooks Brown & Kulik’s original focus, which Curci & Conway (2013: 353) summarise as being “on the phenomenological characteristics of FBMs and emotional factors associated with their formation and maintenance, not on [their] accuracy”. In particular, critical studies have tended to fail to “take account the different involvement that different individuals or social groups have towards a given event” (Curci

& Conway, 2013: 353) and the importance of individual consequentiality to the formation of FBMs (Er, 2003).

Such counter-arguments are persuasive but a recent dedicated review volume on the topic of FBMs (Luminet & Curci, 2008) still highlighted that earlier models tended to overstate the case for the fidelity of memories and that the field overall now recognises many significant limitations, including a number of social and personal factors, that impact the accuracy and enduring nature of flashbulb memories. One article from this volume that directly addressed Modes theory, sought to test it against evidence collected from the Spanish public following the Madrid bombings in 2004. The authors found support for the claim that participation in high arousal 'rituals' improved recall and increased subsequent rumination but they also noted that this effect was superseded in a regression analysis by mass media exposure and concluded that "interpersonal rehearsal seems to be more important than intrapersonal rumination" (Paez et al., 2009: 238). This point echoes arguments raised by Barth (2002) and Houseman (2002) that the role of rehearsal is overlooked in Modes theory's portrayal of imagistic rituals, since previous victims in intense dysphoric rituals often serve as the initiators for following generations. Whitehouse acknowledged the importance of Barth & Houseman's insight (Whitehouse, 2002a, 2004: 112-113) and updated Modes theory to explicitly emphasise the role that *Representational Redescription* (RR) (Karmiloff-Smith, 1992) played in imagistic traditions, through behavioural rehearsal and repetition.

Ignoring the RR amendment to the theory, there are also reasons to question the validity of the specific data presented in Paez et al., (2009), as they use attendance at a political

protest rally to represent a collective imagistic ritual event. Despite efforts by the authors to equate imagistic rituals with political rallies, it is clear that the rally in question lacked almost all of the defining features typically associated with imagistic rituals and even the broader fundamental category of ritual. As such, it offers a poor comparison population to model those engaged in intense personally consequential ritual performances (Whitehouse, 2004). Indeed, Paez et al. (2009) actually specify that they recruited “people not personally affected by the M-11 events”, reinforcing the criticism that researchers tend to ignore the crucial component of personal consequentiality (Curci & Conway, 2013; Er, 2003).

A more relevant case which examined autobiographical accounts at a firewalking ritual event in Greece was recently provided by Xygalatas et al. (2013a: 9). They found that when comparing s of ritual performances to video records, “immediately after the event, participants’ reports revealed limited recall, low confidence and high accuracy” but two months later displayed “more inaccurate memories and higher confidence”. Furthermore, they noted a “spectacular discrepancy between actual heart rates [measured physiologically] and self-reported arousal” (Xygalatas et al., 2013a: 1). They contrast these findings with the high fidelity recall predictions of the Modes theory and alternatively propose a “cognitive resource depletion model”, which suggests that the cognitive effort required for expressive emotional suppression in high ordeal ritual scenarios leads to the formation of inaccurate memories and increases receptivity to received narratives from charismatic authority figures (Schjødt et al., 2013).

Although this example certainly represents a more relevant case than the Spanish political rally, Whitehouse responded by reemphasising that Modes theory is not primarily

focused on whether the episodic memories that imagistic ritual participants develop are accurate but whether they are able to motivate rich personal reflection and strong relational connections with other personalised participants (Whitehouse, 2013b). In line with this Whitehouse argues that imagistic rituals are a poor means of transmitting a standardised doctrine or abstract theology but an effective means of encouraging the transmission of a particular ritual practice. A summary of how the mnemonic predictions of Modes theory in compare with related theories is presented in Table 1.2. Of note is that the theories presented tend to focus on dysphoric rather than euphoric events¹⁷. This reflects a dominant trend in CSR research on ritual prioritising the impact of negative affect, which is potentially due to the broad influence of Modes theory.

The table also reflects an important call for caution issued by Czachesz (2010); namely, that there is evidence for distinctive mnemonic signals from different types of dysphoric affect. In particular, he suggested there was a need to distinguish between rituals with thematic dysphoria and rituals with striking dysphoric visual elements. From reviewing the table, it is clear that some of the predictions of different theories overlap but there are also significant areas of disagreement and uncertainty. Further research to resolve the current uncertainty is necessary but to do so is beyond the scope of the current thesis. Instead, the next section details a different central feature of Modes theory that distinguishes it from some other comparable cognitive frameworks for addressing ritual.

¹⁷ McCauley & Lawson (McCauley & Lawson, 2002; McCauley, 2001) do not distinguish between rituals which are dysphoric and euphoric but instead emphasis that elevated arousal in rituals with 'high sensory pageantry' lead to intuitions about the involvement of a CPS.

Table 1-2: Predictions of the mnemonic effects of dysphoria in rituals

Source of Prediction	Dysphoria in Ritual	
	Improves/Increases	Weakens/Decreases
(Y. I. Russell et al., 2011; Y. I. Russell, Gobet, & Whitehouse, 2014)	+ Attention to detail/precision + Analogical/Relational Thinking	- Ability to “see the big picture”
<i>Modes of Religiosity</i> (Atkinson & Whitehouse, 2011; Richert, Whitehouse, & Stewart, 2005; Whitehouse, 2004)	+ Spontaneous Exegetical Reflection (SER): Volume & Depth of Analogical Thinking + Motivation to transmit ritual + Durable flashbulb/episodic memories + Recall of Ritual Structure + Recall of Participants	- Focus on doctrinal teachings
<i>Rituals with Thematic Dysphoria</i> (Czachesz, 2010)	+ Durable flashbulb memories + Details & circumstances of ritual + Mythical/Moral Narratives	
<i>Rituals with shocking/stressing dysphoric elements</i> (Czachesz, 2010)	+ Memory of shocking dysphoric or stressful elements of a ritual (particularly visual)	- Flashbulb memories - Details & circumstances of ritual - Mythical/ Moral narratives
<i>Ritual Competence Theory: rituals with ‘high sensory pageantry’</i> (McCauley & Lawson, 2002)	+ General attention/awareness that a significant event is occurring + Sensation of involvement of a CPS agent	
<i>Cognitive Resource Depletion</i> (Schjødt et al., 2013; Xygalatas et al., 2013a)	+ Susceptibility to official narratives or charismatic authority + Attention to low level units of behaviour + Confidence in accuracy of recall	- Episodic Memories - Meaningful Individual Narratives

1.6 Cognitively Optimal & Cognitively Costly Religion

A core distinctive feature of Modes theory is that it seeks to account for aspects of religion that are *cognitively costly* rather than processes which appear to be *cognitively optimal* (Whitehouse, 2004: 58). Cognitively optimal accounts of religion focus on aspects of religion that are intuitively appealing to our cognitive systems and are hence attention

grabbing and easily transmitted. MCI theory (J. L. Barrett, 2008; Boyer, 1994, 2000), for instance, represents a cognitively optimal approach as it implies that the violation of a single ontological assumption of intuitive ‘core knowledge’ (Spelke, 2000, 2008) make supernatural concepts appealing and easier to transmit and remember, which makes them prevalent cross-culturally (Boyer & Ramble, 2001; Lisdorf, 2004). Similarly, researchers have found evidence for early developing and cross culturally recurrent tendencies to conceive of invisible beings (J. L. Barrett, 2004; Guthrie, 1993) and recognise a teleological purpose to nature (Casler & Kelemen, 2008; Kelemen & DiYanni, 2005; Kelemen, 1999, 2004). Both of which are linked to humans’ socially orientated minds (Robin I. M. Dunbar, 2003; Lieberman, 2013) and our prolific ability to infer agency in the world¹⁸ (Baron-cohen, 1999; Baron-Cohen, 2000). These characteristics are further argued to make humans receptive to religious ideas (McCauley, 2011), which has led some researchers to claim to suggest that we are ‘Born Believers’¹⁹ (J. L. Barrett, 2012). That is we possess a ‘God instinct’ (Bering, 2010) that promotes dualistic thinking (Bloom, 2004), encourages the endorsement of creationist accounts (Kelemen & DiYanni, 2005), and promotes intuitive beliefs in an afterlife (Bering & Bjorklund, 2004; Bering, 2002). Even those with sophisticated theological beliefs have been demonstrated to utilise more intuitive but theologically incorrect versions when placed under cognitive stress (Barrett, 1999).

There are also cognitively optimal accounts of ritual including ritual competence theory (Lawson & McCauley, 1990; McCauley & Lawson, 2002), which emphasises the importance

¹⁸ Barrett has even suggested that we possess a *hyperactive* agency detection device (HADD) that readily interprets ambiguous stimuli as signals of agency (J. L. Barrett, 2004) .

¹⁹ Though see Jong, Kavanagh, & Visala (2015) for why ‘Born Idolaters’ would be a better description.

of *Theory of Mind* (ToM) (Baron-cohen, 1999; Frith & Frith, 2005) and a universal *causal action representation system* in determining the structure and expected outcomes of rituals. Boyer & Liénard's (Boyer & Liénard, 2006; Liénard & Boyer, 2006) 'Hazard Precaution' model also represents cultural rituals as "parasitic" by-products of more fundamental hazard-detection and precaution systems (Boyer & Liénard, 2006: 15). Although, it should be noted that Boyer & Liénard consciously distance their approach from offering a theory of ritual and instead focus on explaining ritualized behaviour. This means that a focus on low level cognitive processes is somewhat inevitable but they do also recognise that ritualized behaviour is constrained by "high-level cognitive control" (Boyer & Liénard, 2006: 12) and thus does not entirely rest on non-reflective cognitive processes.

The common element unifying both of these theories is that they present rituals as being able to capture attention and motivate performance primarily due to their ability to stimulate universal cognitive processes which are fast, automatic and unconscious (Evans, 2008; Kahneman, 2011) and whose proper domain relates to an unrelated evolutionarily adaptive value. Responding appropriately to hazards (Boyer & Liénard, 2006), detecting unseen threats (J. L. Barrett, 2004; Guthrie, 1993), and building reliable causal models (McCauley & Lawson, 2002) are all important cognitive abilities for a vulnerable primate species, and in cognitive optimal accounts these biases are also used to account for the widespread fascination with rituals. In such accounts, the complex exegetical traditions that typically surround ritual performance are regarded as largely inconsequential: "while the meanings associated with rituals may vary... for some explanatory purposes meanings hardly matter at all" (McCauley & Lawson, 2002: 10).

This represents a key departure with Modes theory, as alternatively Whitehouse argues that although the activation of universal cognitive mechanisms can account for certain culturally recurrent features of religion, like the prevalence of MCI gods, these explanations fail to adequately address the “tendency for religious traditions to eschew... cognitively optimal concepts and to peddle more complex bodies of knowledge in their stead” (Whitehouse, 2004: 46). In contrast to cognitively optimal approaches, Modes theory contends that doctrinal and imagistic rituals represent two successful but divergent strategies for memorising and transmitting complex symbolic and exegetical knowledge (Atkinson & Whitehouse, 2011). That is, the repetition and routinization of the doctrinal mode enable large, complex networks of religious knowledge “to be stored in semantic memory, reproduced stably and spread efficiently as oral tradition” (Atkinson & Whitehouse, 2010: 51). While the vivid episodic experiences of the imagistic mode can trigger the build-up of internally generated insight, which is then later strengthened and reinterpreted through interactions with the community and the use of analogic principles (Atkinson & Whitehouse, 2010: 51). In Modes theory then the doctrines and interpretations that surround rituals are not an irrelevance but a crucial component that effects transmission and other broader characteristics of groups.

Modes theory is not the only approach to recognise that it is insufficient to focus solely on the role of innate intuitions when addressing the transmission of religions. Henrich & Gervais (2010), for example, also point out that there are a number of counterintuitive figures, such as Mickey Mouse and Zeus, which meet the stated criteria for success laid out in cognitively optimal accounts but yet they do not inspire the predicted levels of belief in

modern societies. Hence, Henrich (2009) emphasises the need to consider other contextual factors involved in religious transmission, including the importance of costly ritual practices and their ability to serve as ‘credibility enhancing displays’ (CREDs) (Henrich, 2009; Lanman, 2012; Wildman & Sosis, 2011). This interpretation is favoured by researchers who contend that religious belief is a naturally selected adaptation for cooperative group living (Alcorta & Sosis, 2005; Atran & Henrich, 2010; Bering, McLeod, & Shackelford, 2005; Bulbulia, 2007; Bulbulia et al., 2013; Henrich, 2009; Norenzayan, 2013; Purzycki et al., 2014; Wildman & Sosis, 2011; Wilson, 2010).

There is an on-going debate about the evolutionary role of religion within the CSR field (Dow, 2008; Schloss, Barrett, & Murray, 2010; Sosis, 2009) but the dominant position at present is that it probably plays some role in cultural group selection (Bulbulia et al., 2013). Advocates of this position hold different positions on the precise mechanisms involved but collectively they recognise the combined importance of cognitively optimal and costly aspects of religion. This means that the insights provided by cognitively optimal approaches are not regarded as incompatible, but rather complementary with broader evolutionary models, such as Modes theory (Gervais, Willard, Norenzayan, & Henrich, 2011; Whitehouse, 2004). Furthermore, although there are competing models that address the cognitively costly aspects (Schjødt et al., 2013; Whitehouse, 1996; Xygalatas et al., 2013a) and the evolutionary history of religion (Baumard, Hyafil, Morris, & Boyer, 2014; Norenzayan, 2013; Weeden & Kurzban, 2013) various aspects of these rival theories can potentially be synthesised productively.

To illustrate this, it is instructive to consider Modes theory and costly signalling theories both of which discuss costly rituals and appear to provide distinct explanations of their social and psychological effects. However, when examined carefully it is clear that the theories are not mutually exclusive and in fact have a very different emphasis. In Modes theory the focus is primarily on the performers of dysphoric rituals and the impact that participation has on their autobiographical memory and the bonds they form with other performers. Whereas with costly signalling theories the focus is primarily on the observers of ritual performances and broader community levels of trust. The two theories should thus not be regarded as necessarily antagonistic but rather considered as potentially complementary. This does not mean that there are never conflicting predictions but rather that care must be taken to ensure that contrasted theories really are addressing the same issue at the same level. If given due consideration, there are likely many debates within the theoretical landscape of the CSR field that would become less polarising.

1.7 Identity Fusion & Group Bonding

One of the most significant distinctions that the Modes theory posits between doctrinal and imagistic modes is in the different types of bonds they involve (Whitehouse & Lanman, 2014; Whitehouse, 2013c). In the doctrinal mode, the bonds created between ritual members are diffuse and extend across a large 'imagined community' (B. Anderson, 1983), whereas in the imagistic mode the bonds are intensely localised and directed at specific individuals who share specific ritual experiences. The distinction between these two types of bonds relates closely to a distinction between categorical and relational bonds proposed by Brewer & Gardner (1996). Relational bonds involve a personal connection with particular

people and are characteristic of the kind of bonds experienced in families. In contrast, categorical bonds entail a more abstract connection based on shared characteristics or symbolic identity markers, such as these found with employees working for the same company. Brewer & Gardner's (1996) formulation of these different ties was intended as a corrective amendment to the 'social identity perspective', the dominant theoretical model for explaining group behaviour in social psychology (Abrams & Hogg, 1988; Hogg, 2006; Hornsey, 2008; Tajfel & Turner, 1985; J. C. Turner & Reynolds, 2010; J. C. Turner, 1999).

The social identity perspective is the primary lens through which psychologists have viewed group behaviour and group bonding and is comprised of 'social identity theory' (SIT) (Tajfel & Turner, 1979) and 'self-categorisation theory' (J. C. Turner, Hogg, Oakes, Reicher, & Wetherell, 1987). The central insight from these theories is that there is an identifiable and dialectical distinction between the personal and the social self. The personal self refers to those aspects of identity that relate to an individual and their singular personality (funny, aggressive, shy), whereas the social self refers to aspects of identity associated with membership of specific groups (Irish, Muslim, Liberal). Individuals are not envisioned as having a single 'social self', instead they possess a series of overlapping social categories which are made salient in different circumstances. SIT also posits a principle of 'functional antagonism' between the social and personal self. This principle means that in scenarios where social identity is made salient, personal identity will fade in importance and vice versa. One implication of this is that, since personal identity will recede into the background in group scenarios, any pro-social behaviour directed towards an in-group, or hostility directed towards an out-group, is typically motivated by the relevant *social* identity

activated in that context (Hopkins et al., 2007; Levine & Crowther, 2008). This extinguishing of the personal self in group scenarios is argued to lead to the 'depersonalization of self-representation' (Rosenberg, 1987), in which individuals "come to perceive themselves more as the interchangeable exemplars of a social category than as unique personalities defined by their individual differences from others" (J. C. Turner et al., 1987: 50). The stronger the identification with a particular group the stronger this depersonalization is experienced. And this depersonalization occurs not just at a personal level but extends to other members of the relevant in-group and out-groups, who are also perceived as largely interchangeable (Hogg, 2006: 118).

There is a substantial experimental literature that attests to the validity of SIT (for recent reviews see: Hornsey 2008; Turner & Reynolds 2010; Stryker & Burke 2000) but a core pillar is derived from studies using the so called 'minimal group paradigm' (MGP) promoted by Henri Tajfel and colleagues during the 1970s and 1980s. In these studies, it was demonstrated that even when allocated to groups on the basis of meaningless and arbitrary criteria, such as a coin flip, individuals would almost immediately demonstrate an ingroup bias in intergroup tasks. And in some cases were even willing to allocate a lower total overall reward for their group, if it led to a greater disparity in rewards when compared with the opposing group (Tajfel et al., 1971a; Tajfel, Billig, Bundy, & Flament, 1971b; Tajfel & Turner, 1985; Tajfel, 1970). Tajfel argued that these results not only demonstrated the fundamental nature of group cognition processes but also illustrated how interpersonal relations with individuals are immediately transformed when group categorisation is invoked to become focused on intergroup categorical interactions (Tajfel & Turner, 1985).

The MGP literature is too vast to review here but in summation the early experiments were followed by replications of the main effects and studies which sought to establish boundary conditions and rule out alternative accounts (Claire & Turner, 1982; Moghaddam & Stringer, 1986; Oakes & Turner, 1980; M. J. Platow, McClintock, & Liebrand, 1990). But there was also a growth of critical studies that questioned the consistency of the ingroup bias effect (Bourhis, Sachdev, & Gagnon, 1994; Diehl, 1990; Rabbie & Horwitz, 1988) and argued that there was important cultural variation (Falk, Heine, & Takemura, 2013; Gaertner & Insko, 2001), with particular attention being paid to the importance of perceptions of interdependence and self-interest (Gagnon & Bourhis, 1996; Jin, Yamagishi, & Kiyonari, 1996; Karp, Jin, Yamagishi, & Shinotsuka, 1993).

The focus on the MGP in SIT research was also a topic of criticism (Bornstein et al., 1983; Schiffmann & Wicklund, 1992) and as a result from the early 1990s efforts were made to design experiments that incorporated existing social categories and provided greater recognition of the complexity of intragroup dynamics. Hence, despite tendencies to perceive the in-group as homogenous (Haslam, Oakes, Turner, & McGarty, 1996; Hogg & Turner, 1987; Simon & Hamilton, 1994), recent research has also indicated that highly identified group members do recognise internal differences between ingroup members, largely through the degree to which members resemble an ideal group prototype. Differing degrees of perceived proto-typicality are associated with differential roles and varying influence within the group; higher prototype resemblance increases status and influence and lower resemblance results in marginalisation (Cheng, Fielding, & Terry, 2011; Hogg, 2000; Van Kleef, Steinel, Van Knippenberg, Hogg, & Svensson, 2007). However, even with

this amendment the types of bonds that are being described in SIT are still of the collective and categorical variety. The theory is thus ill suited to account for relational bonds, such as those associated with imagistic ritual groups. This limitation is what led Brewer and Gardner (1996) to argue for the importance of recognising the key role of relational bonds and the potential for groups to form based around such bonds.

In terms of measurement of the categorical bonds described in SIT there is a plethora of social/group identification scales (for a recent detailed review see Leach et al. 2008). Again it is beyond the scope of this thesis to delve deeply into the various approaches taken but a brief sketch of three popular group identification measures is provided for illustrative purposes and because variations of these scales are employed in the studies reported in Chapters 2 & 3. The first measure to be discussed is Mael and Ashforth's popular six item measure of social identification with organizations (Fig 1.4).

Figure 1-4: Verbal measure of 'organizational identification'

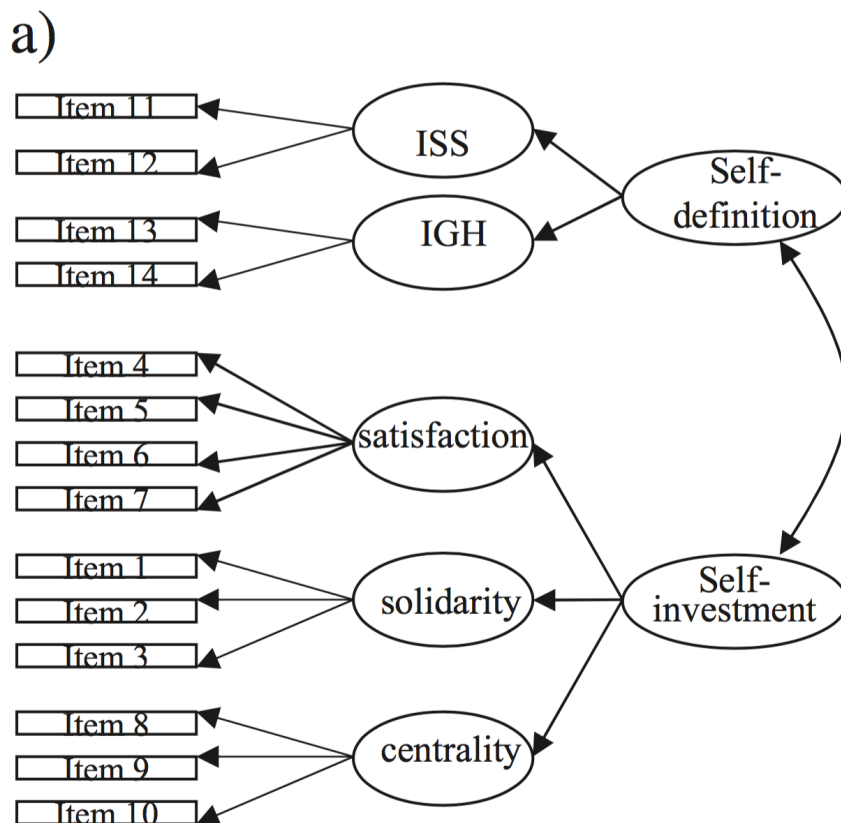
1. When someone criticizes (name of organization), it feels like a personal insult.
2. I am very interested in what others think about (name of organization).
3. When I talk about this organization, I usually say 'we' rather than 'they'
4. This organization's successes are my successes
5. When someone praises this organization, it feels like a personal compliment.
6. If a story in the media criticized the organization, I would feel embarrassed.

Source: (Mael & Ashforth, 1992)

The items in this scale reflect the key aspects of the SIT paradigm, specifically that personal identity is replaced with categorical identification and an intergroup orientation is emphasized. This scale is popular due to its short concise format and the original publication that the scale appeared in has been cited over 7,000 times (Ashforth & Mael,

1989). But the Mael & Ashforth scale is not by any means the accepted standard measure in the relevant fields of social psychology, instead as the recent review paper by Leach et al. (2008) argues there has been “little agreement about how identification should be conceptualized or measured”. Leach et al. (2008) attempted to resolve this uncertainty by synthesising popular previous measurement scales into a novel fourteen item scale (Fig.1.6) based on a five component, two dimensional structure which they demonstrated over seven studies to have structural, predictive, and discriminant validity (Fig 1.5).

Figure 1-5: Component structure of in-group identification



Note: ISS = Individual Self-Stereotyping, IGH= In-Group Homogeneity
Source: (Leach et al. 2008)

Although the Leach et al. (2008) measure illustrates the complexity involved with examining identification processes, another recent paper by Postmes et al. (2013) has

conversely argued that a single item self-identification measure may be sufficient in certain cases. Postmes et al. (2013) simply asked participants to respond on a seven-point scale to the statement: “I identify with [my group]” and found that the measure displayed good reliability and evidence of convergent, divergent, and test-retest validity even when compared against longer scales, such as the Leach et al. (2008) measure. The main point in referencing this study is not to deny that there are multiple components to group identification but to argue that fundamentally it is a “sufficiently homogenous” construct that it can be operationalized with a single item.

Figure 1-6: Verbal measure of ‘in-group identification’

Items Measuring In-Group Identification	
(Group-Level) Self-Investment	(Group-Level) Self-Definition
<i>Solidarity</i>	<i>Individual Self-Stereotyping</i>
1. I feel a bond with [In-group]. (Adapted from Cameron, 2004; Doosje et al., 1998.)	11. I have a lot in common with the average [In-group] person. (Adapted from Spears et al., 1997.)
2. I feel solidarity with [In-group].	12. I am similar to the average [In-group] person. (Adapted from Doosje et al., 1995; Spears et al., 1997.)
3. I feel committed to [In-group]. (Doosje et al., 1995)	<i>In-Group Homogeneity</i>
	13. [In-group] people have a lot in common with each other. (Adapted from Spears et al., 1997.)
<i>Satisfaction</i>	14. [In-group] people are very similar to each other. (Adapted from Ellemers, Kortekaas, & Ouwerkerk, 1999; Spears et al., 1997.)
4. I am glad to be [In-group]. (Adapted from Cameron, 2004; Doosje et al., 1998; Luhtanen & Crocker, 1992.)	<i>Excluded Items</i>
5. I think that [In-group] have a lot to be proud of. (Ellemers, Kortekaas, & Ouwerkerk, 1999.)	15. I see myself as [In-group]. (Doosje et al., 1998.)
6. It is pleasant to be [In-group]. (Doosje et al., 1998.)	16. I identify with other [In-group] people. (Doosje et al. (1998); Ellemers, Kortekaas, & Ouwerkerk, 1999.)
7. Being [In-group] gives me a good feeling. (Adapted from Cameron, 2004; Luhtanen & Crocker, 1992.)	17. Being [In-group] just feels natural to me. (Doosje et al., 1998.)
<i>Centrality</i>	18. I feel (personally) implicated when [In-group] people are criticized. (Adapted from Jackson, 2002.)
8. I often think about the fact that I am [In-group]. (Adapted from Cameron, 2004.)	19. [In-group] are an important group to me. (Doosje et al., 1998.)
9. The fact that I am [In-group] is an important part of my identity. (Adapted from Luhtanen & Crocker, 1992.)	20. I have a lot of respect for [In-group] people. (Ellemers, Kortekaas, & Ouwerkerk, 1999.)
10. Being [In-group] is an important part of how I see myself. (Adapted from Doosje et al., 1998; Ellemers, Kortekaas, & Ouwerkerk, 1999; Luhtanen & Crocker, 1992.)	

Received May 15, 2006
Revision received October 9, 2007
Accepted October 15, 2007 ■

Source: (Leach et al. 2008: 165)

None of the measures discussed above address the fundamental shortcomings of the traditional SIT model in regards to the lack of emphasis on relational bonds. But an alternative and newly developed theory termed ‘identity fusion’²⁰ is claiming to have resolved this oversight and identified an important, previously ignored, form of social bonding (Gómez, Brooks, et al., 2011; Swann, Gómez, Dovidio, Hart, & Jetten, 2010; Swann, Gómez, Huici, et al., 2010; Swann et al., 2014, 2009, 2012). Fusion is defined by a feeling of co-essence with a group and a sense of reciprocal agency. Rather than self-identity giving way to a group identity, as posited with group identification in SIT, fusion entails a *synthesis* of self and group identities. In a fused state an individual’s concept of self and their personal agency remain firmly intact but they also experience a deep connection to their group:

“Fusion is marked by a visceral feeling of oneness with the group. This feeling is associated with a highly agentic-personal-self and unusually porous, highly permeable borders between the personal and social self... In addition, the porous borders will encourage highly fused persons to develop strong relational bonds with their fellow group members and maintain lasting commitment to the group” (Swann et al., 2012).

Fusion is proposed to come in two varieties *local* and *extended*. *Local* fusion is when “group members form relational bonds with others whom they have direct personal contact”, and *extended* fusion is when people “project the relational bonds normally associated with local fusion onto large groups”, such as a nation (Swann et al., 2012: 4). The process by which relational bonds can be projected onto a collective group of strangers

²⁰ ‘Identity fusion’ is primarily referred to throughout the rest of this thesis as ‘fusion’.

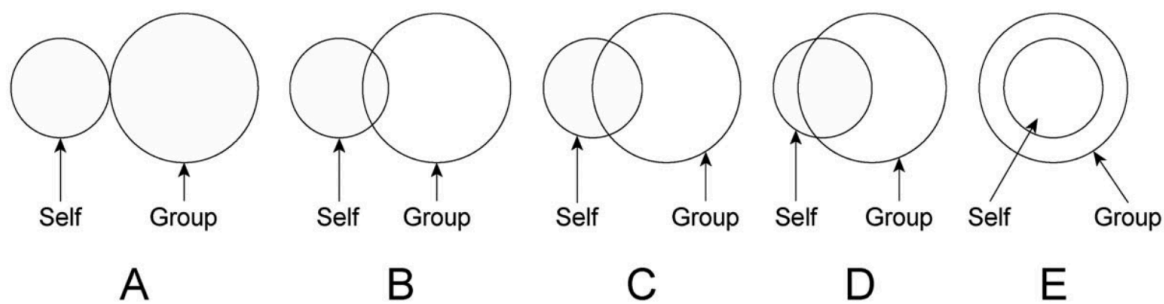
remains poorly understood but preliminary research, conducted with cross cultural samples from China, India, the US, and Spain, suggest that it is related to the perception of sharing core characteristics, both biological (genes) and psychological (core values), with other group members (Swann et al., 2014). When these shared characteristics are salient this can result in the creation of a fictitious kin group (Atran & Henrich 2010; Swann et al. 2012; Whitehouse et al. *submitted*).

Fusion was originally measured pictorially (Fig. 1.7) and considered to be a dichotomous state, but the development of a more nuanced verbal scale (Fig. 1.8) led to the discovery that it was actually distributed continuously, meaning that there were varying degrees of fusion²¹. Following this, a large number of studies, including over 30 individual experiments and surveys, have validated cross culturally that fusion measures a psychological state distinct from group identification, resulting in differential outcomes on a host of outcomes (for reviews see Swann & Buhrmester 2015; Swann et al. 2012). In particular, there is strong evidence that fusion is a stronger predictor than group identification of individuals' willingness to endorse fighting or sacrificing themselves for their group (Swann, Gómez, Dovidio, et al., 2010; Swann et al., 2014, 2009). Such sacrifices can be orientated towards intergroup conflict but also include broader costly pro-group behaviour, such as donating personal funds to needy group members (Semnani-Azad, Sycara, & Lewis, 2012; Swann, Gómez, Huici, et al., 2010), performing better in a tiring computer game (Swann, Gómez, Huici, et al., 2010), and whistleblowing on cheaters

²¹ The recently developed dynamic identity fusion index (DIFI) provides another visual measure of fusion that can capture a much greater level of nuance to responses (Jimenez et al., 2015).

(Buhrmester, 2013). Furthermore, highly fused individuals have been found to intensify their loyalty in response to ostracism, in contrast to highly identified individuals who did not display such tendencies (Gómez, Morales, Hart, Vázquez, & Swann, 2011). Highly fused individuals have also been shown to perceive their personal fortunes as strongly associated with both the success and the failures of their groups, whereas individuals with high identification only saw their personal circumstances as linked to group successes (Buhrmester et al., 2012).

Figure 1-7: Pictorial measure of identity fusion



Source: (Swann et al., 2009)

Figure 1-8: Verbal measure of identity fusion

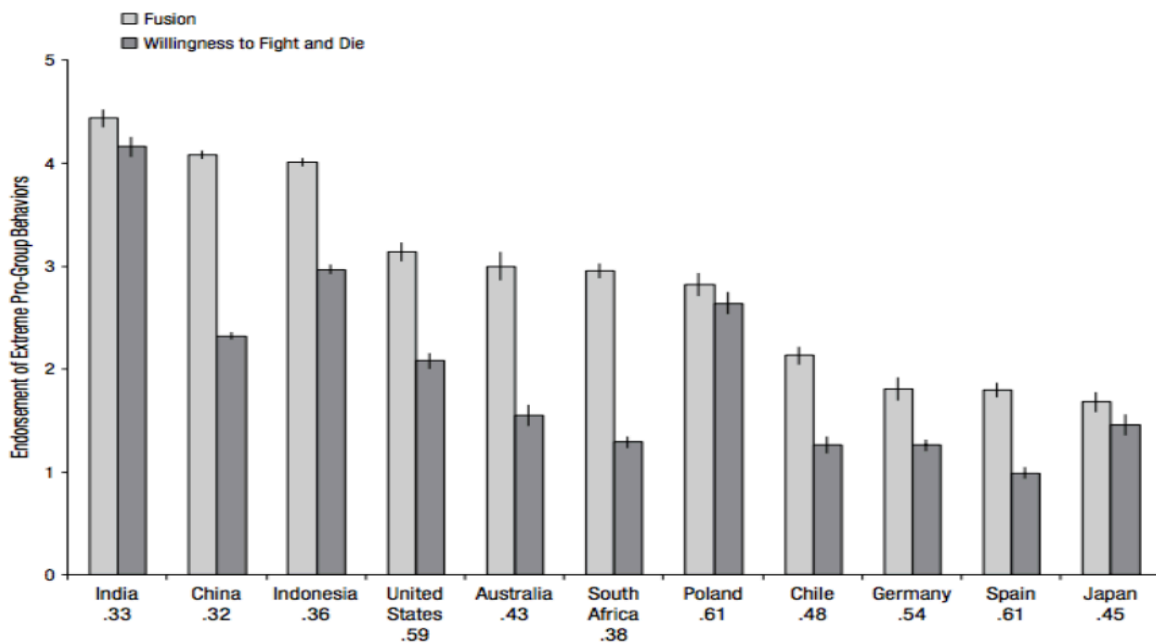
1. I am one with my country.
2. I feel immersed in my country.
3. I have a deep emotional bond with my country.
4. My country is me.
5. I'll do for my country more than any of the other group members would do.
6. I am strong because of my country.
7. I make my country strong.

Source: (Gómez, Brooks, et al., 2011)

Several factors that mediate the relationship between costly pro-group actions and identity fusion have been identified including: the perception of familial ties (Buhrmester,

Fraser, Lanman, Whitehouse, & Swann, 2014; Swann et al., 2014; Whitehouse, McQuinn, Buhrmester, & Swann, 2014), shared belief in sacred values (Atran et al. 2014a; Atran et al. 2014b), physiological and emotional arousal (Swann, Gómez, Huici, et al., 2010; Swann et al., 2014), feelings of invulnerability (Gómez, Brooks, et al., 2011), and group directed agency (Gómez, Brooks, et al., 2011).

Figure 1-9: Correlations between rates of fusion and willingness to fight or die



Source: (Swann & Buhrmester, 2015; Swann et al., 2014)

Importantly, although the majority of fusion studies have focused on student populations and primarily presented scenarios of hypothetical sacrifice for the nation state, more recent studies have examined fusion in more ecologically valid settings. Whitehouse et al. (2014), for instance, have examined fusion amongst Libyan revolutionary fighters in the 2011 civil war, and Atran et al. (2014a, 2014b) surveyed levels of fusion and sacred values in two Moroccan communities associated with support for militant jihad. Outside of an intergroup conflict scenario, Swann et al. (2015) recently conducted a study with Spanish

transsexuals and found that fusion with gender was able to effectively predict later willingness to perform genital sex reassignment surgery.

The significance of this new flurry of research for Modes theory is that fusion represents a plausible psychological mechanism for the bonds formed during imagistic rituals (Whitehouse & Lanman, 2014). Specifically, imagistic bonds are presented in Modes theory as being of an intense and highly relational nature; they are directed at specific ritual co-participants and are based on the sense of having shared an important experience. The personalised and relational nature of such ties seems to directly contradict the predictions of SIT but they fit well with the fusion theory, particularly with the concept of local fusion. A potential connection between imagistic ritual events and fusion was first discussed in Swann et al. (2012) but more recently was elaborated into a “testable theory of the relationship between ritual, cohesion, and cooperation that more precisely connects particular elements of ‘ritual’... to two particular forms of ‘social cohesion’: group identification and identity fusion” (Whitehouse & Lanman, 2014).

Building on the distinctions discovered between group identification and identity fusion, this new theoretical model argues that the processes represent two different varieties of ‘social glue’ (Whitehouse, 2013c, 2015). In this model, identity fusion, and the extreme willingness to sacrifice it promotes, relies on the creation of psychological kinship. Psychological kinship is derived from our evolved kin recognition heuristics (Bressan & Kramer, 2015), which operate through perceptions of phenotypic similarity (DeBruine, 2005; Krupp, DeBruine, Jones, & Lalumiere, 2012), but also, crucially, through socially derived information, such as “the perception that one shares with others episodic memories

that are essential components of one's autobiographical self-concept" (Whitehouse & Lanman, 2014: 677).

Imagistic rituals, particularly those with highly dysphoric elements, are argued to represent one highly effective way of establishing rich, personally consequential autobiographical memories that create deep links with other ritual participants. Some initial evidence in support of this connection comes from recent research on rebel paramilitary groups, which have been found to adhere to the ritual dynamics posited by Modes theory (Whitehouse & McQuinn, 2012). Research conducted with revolutionary fighters during the 2011 civil war in Libya found that nearly half of frontline combat troops reported being more strongly fused with their battalion than they were with their biological family (Whitehouse et al., 2014). Similarly, a study of Vietnam veterans found that those who had experienced particularly intense dysphoric events, such as witnessing comrades being killed or wounded, were more fused with the identity "Vietnam veterans" but not more identified (Matthews, Lanman, Swann, & Whitehouse, *in prep.*).

These findings correspond with growing evidence that perceptions of familial ties mediate the relationship between fusion and sacrificial behaviour. Buhrmester et al. (2014: 253), for instance, demonstrated that, after the Boston Marathon bombing in 2013, the "degree to which participants reported perceiving fellow Americans as psychological kin... mediated links between fusion... and [the provision of] various forms of support to the bombing victims". Imagistic rituals are also associated with higher physiological arousal and another recent study (Swann, Gómez, Huici, et al., 2010) demonstrated that highly fused individuals displayed significantly greater pro-group behaviour following exercise.

These preliminary findings suggest that high arousal imagistic rituals, particularly dysphoric rituals, may serve as an effective source for generating identity fusion. However, a plausible alternative account is that dysphoric imagistic rituals are typically only performed by individuals who are already highly fused with their ritual group. The causal relationships require further investigation and experimental examination to help unpack.

In contrast to the intense relational bonds of fusion, the second proposed variety of 'social glue' is the categorical bonds generated by the group identification processes of coalitional psychology. The key distinction here is that group identification enables the bonding of "larger assemblages of more distantly related individuals" (Whitehouse & Lanman, 2014: 678) as the relationship is with an abstract social category as opposed to individual group members. A paradigmatic example of this effect is found in minimal group studies, which demonstrate that even arbitrary group categories amongst strangers can prompt in-group bias and out-group discrimination (Diehl, 1990; Schiffmann & Wicklund, 1992; Tajfel & Turner, 1979).

In relation to Modes theory it is argued that the routinized rituals of the doctrinal mode serve to enhance identification with a group, by providing repeated reinforcement of beliefs and practices that serve as group identity markers in semantic memory. Doctrinal traditions thus do not rely on "activating lasting episodic memories and private reflection but rather through... activating procedural and semantic memory for conventional practices shared by communities too large for their members to know each other personally" (Whitehouse & Lanman, 2014: 683). One issue that identification with large collectives pose is that it can make individuals and groups vulnerable to exploitation from

‘free riders’. Free riders are those who utilize group identity markers in order to gain access to group benefits and cooperation but fail to contribute to the group or reciprocate to other members. It has been suggested that because of potential exploitation from free riders in groups where the doctrinal Mode is dominant, there is a need for additional mechanisms to promote cooperation, such as morally concerned punitive deities (Bering, 2010; Norenzayan, 2013), or credibility enhancing displays (Bulbulia & Sosis, 2011; Henrich, 2009). This explanation might also help to account for the occurrence of imagistic practices within doctrinal traditions.

Placed within an evolutionary framework, the imagistic mode and the intensely relational bonds of identity fusion are envisioned as having developed earlier and enabled small scale societies to overcome collective action problems, particularly as those that involve high risks also provide strong temptations to defect, such as in warfare or big game hunting (Whitehouse & Lanman, 2014: 681-682). A less personalised form of bonding that could generate commitment to larger groups became necessary with the growth of larger societies with more complex divisions of labour. In the Modes framework this is supplied by later developing, group identification processes which are argued to accompany the growth of the doctrinal mode of religion. This account contradicts the perspective offered by researchers such as Robin Dunbar (2009) who argues that it was large group living which generated the selective pressures on mechanisms that promote tight sub-group coalitional alliances, in order to buffer individuals in such societies against threat and attack.

The precise temporal and causal pathways involved, including the factors that enabled doctrinal religions to emerge (Baumard et al., 2014) remain speculative, but there is emerging archaeological research that suggests that the transition from small-group living to larger, more complex societies was accompanied by a change in the frequency and emotionality of ritual (Whitehouse & Hodder, 2010).

In summary, the Modes of Religiosity theory has been substantially refined and updated from its earlier iterations and more recent versions emphasise how imagistic and doctrinal religions relate not just to distinctions in semantic and episodic memory but also to distinct psychological affiliative processes (Swann et al., 2012; Whitehouse & Lanman, 2014). The theory has also expanded in recent years to address group dynamics in contexts outside of religious traditions (Whitehouse et al., 2014; Whitehouse & McQuinn, 2012). On the one hand, this is to be expected as ritual practices are clearly not restricted to religious contexts and yet the core association between religion and ritual remains well established (Purzycki & Sosis, 2013; Rappaport, 1999). To what extent ritual practices can determine the features of non-religious communities as Modes theory predicts remains an open question. Furthermore, there is a clear need to test the links proposed between imagistic rituals and identity fusion and doctrinal rituals and group identification. Fortunately, such questions are amenable to empirical investigation and this thesis seeks to contribute to ongoing efforts to clarify the connections by exploring the proposed relationships between imagistic rituals and identity fusion processes, addressing the role of collective affective arousal in pro-group behaviour.

1.8 Affective Experiences & Group Bonding

Drawing on pioneering efforts (Darwin, 1965; James, 1884), a vast psychological literature has developed devoted to exploring the topic of emotion. In particular, there has been significant research on how best to categorise emotional experiences (Bradley & Lang, 1994; Ekman, 1992; Larsen & Diener, 1992; J. A. Russell, 1980; Tellegen, Watson, & Clark, 1999; Watson, Clark, & Tellegen, 1988), whether emotional categories are consistent cross cultural (Ekman, 1992; Elfenbein & Ambady, 2002; Ortony & Turner, 1990; J. A. Russell, 1991; Sauter, Eisner, Ekman, & Scott, 2010; Wierzbicka, 1986), and what the social and evolutionary function of emotions are (Damasio, 2008; Haidt, 2001; MacLean, 1993; Shariff & Tracy, 2011; Spoor & Kelly, 2004).

The anthropological literature also explores the issue of emotion (A. Beatty, 2005, 2014; S. E. Beatty, Yoon, Grunert, & Helgeson, 1996; Lutz & White, 1986), focusing primarily on how certain emotions are understood and expressed (differently) in given cultures (M. Z. Rosaldo, 1980; R. Rosaldo, 1980; Shweder, 1993). Unfortunately, in the anthropological literature there is very limited agreement on terminology and an appreciable lack of enthusiasm for systemisation. Beatty, for instance, highlights two common objections to any attempts to generalise about emotion: the first is that they “are *first-person experiences* and not reducible to any of their ingredients”, and the second, is that “emotions are biographical: primed by evolution, to be sure; shaped by culture; constrained by subject position; but given personal relevance and intensity by individual history” (Beatty 2014: 551-552). The solution proffered by Beatty is to advocate a narrative approach, which is stated to be the *only* approach that “can capture both the particularity and the temporal

dimension of emotion, restoring verisimilitude and fidelity to experience” (2014: 545). This encapsulates the dominant methodological trend of mainstream anthropology and illustrates why the psychological literature on emotion is more relevant for efforts that seek to empirically investigate the topic. It is beyond the scope of this thesis to explore the debates in anthropology in depth, or to review the vast psychological literature on emotion in general. Instead, this section focuses on providing an outline, and justification for, the structural model of emotion employed throughout the rest of this thesis and summarising existing research on how collective emotional states affect group bonding.

One group of psychological theories of emotion focuses on identifying *basic emotions* that are posited to be universally recurrent and discrete, in the sense of involving distinctive facial expressions and biological processes (Ekman, 1992; Ekman et al., 1987). The most prominent advocate of this approach, Paul Ekman, has identified seven basic emotions which have good evidence for universality²²: anger, disgust, fear, happiness, sadness, and contempt (Ekman & Heider, 1988; Ekman, 1992; Ekman et al., 1987), and, at least, a further ten likely candidates (Ekman & Cordaro, 2011). Although such theories of basic emotion have proven extremely influential and have a substantial number of studies supporting their central claims (Ekman, 1994; Elfenbein & Ambady, 2002; Izard, 1994, 2007; Panksepp, 1998, 2007; Sauter et al., 2010), they are also the subject of substantial criticism.

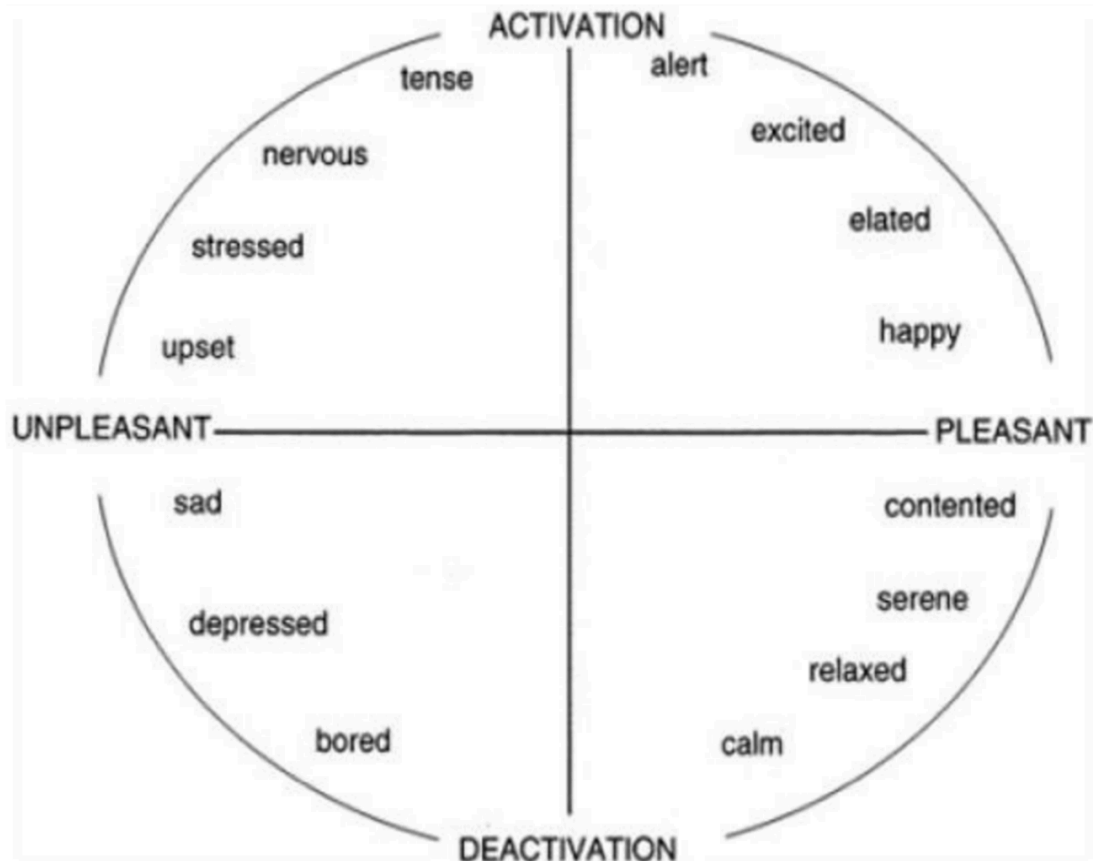
²² Ekman actually varies across articles in whether each of these emotions has met the evidential requirement to be considered basic. For instance, contempt is often excluded from earlier summaries but Ekman repeatedly argued for its inclusion and it was added to later lists. He has also continued to develop his list of characteristics for an emotion to be considered basic with the most recent criteria listing 13 required characteristics (Ekman & Cordaro, 2011).

Hence, in a series of meta-analyses of studies correlating autonomic activity and affect, Cacioppo et al. (2000: 184) found “that even a limited set of discrete emotions... cannot be fully differentiated by visceral activity alone”. Similarly, in a wide ranging critical review, Russell (1994: 102) highlighted a variety of methodological problems and argued that when better controlled studies were conducted the results were “less supportive or non-supportive” of Ekman’s position. This conclusion was echoed a decade later in another detailed review by Barrett (2006), which concluded that when considered in totality there was a tremendous amount of inconsistent evidence. Barrett’s conclusion was that on balance, despite the existence of some positive studies, the hypothesis should be discarded (Barrett et al. 2007).

The debate around this topic is unlikely to be resolved anytime soon but there is an alternative theoretical model that has proven equally influential; the circumplex model of affect (Larsen & Diener, 1992; J. A. Russell, 1980, 2003). *Affect* is an umbrella term that relates to both positive and negative experiences that can vary in terms of strength, duration, and specificity. This makes the term useful for addressing moods (less intense, more diffuse) and emotional episodes (more intense, specific). The circumplex model is derived from a dimensional approach that seeks to represent the dominant intercorrelations found in research on emotional experiences (Bradley & Lang, 1994; Mehrabian, 1995; J. A. Russell, 1980; Watson et al., 1988). Such an analysis yields a model that proposes that all emotional experiences arise from two fundamental neurophysical systems: one associated with valence (pleasure to displeasure) and the other with arousal/activation (alert to calm). In experiential terms these two dimensions combine into

a singular feeling, referred to as *core affect*, that is constant and consciously accessible but usually non-reflectively experienced, as with body temperature (J. A. Russell & Barrett, 1999; J. A. Russell, 2003, 2009). Emotional episodes rely on these same underlying systems, but also draw on culturally constructed categories to label experiences.

Figure 1-10: The circumplex model of affect



Source: (Posner, Russell & Peterson 2005)

The crucial distinction between this model and that of basic emotion theories is that emotional states are not regarded as discrete entities but rather as constellations of interrelated factors that derive from fundamental shared neurophysical systems. Russell argues that because of their shared aspects the circumplex model can accurately represent any affective state, broadly categorising them into four quadrants (J. A. Russell, Weiss, &

Mendelsohn, 1989; J. A. Russell, 1980) (Fig.1.10). Under this model fear would be categorised as having negative valence and high activation; depression as negative valence with low activation; elation as positive valence and high activation and relaxation as positive valence with low activation.

The circumplex model is supported by substantial converging evidence from psychometric studies of emotion (Bradley & Lang, 1994, 1999; Crawford & Henry, 2004; P. J. Lang, Bradley, & Cuthbert, 2008; J. A. Russell & Mehrabian, 1977; Watson et al., 1988) and preliminary supportive evidence from cognitive neuroscience and neuroimaging studies (J T Cacioppo, Petty, Losch, & Kim, 1986; John T Cacioppo et al., 2000; Posner, Russell, & Peterson, 2005). However, the neurophysical connections postulated between the affective and arousal systems remain highly speculative and the two factor model also reduces the importance of a third factor- dominance, which is posited as a core element in Pleasure-Arousal-Dominance models (PAD) (Bradley & Lang, 1994; Mehrabian & Russell, 1974; J. A. Russell & Mehrabian, 1977). Dominance in this model represents the extent to which one feels in control/dominant as opposed to submissive during an emotional experience. Hence, although fear and anger can both be defined as high arousal unpleasant emotions, fear tends to be an emotion associated with submission, while anger is more strongly associated with expressions dominance. However, for the purposes of the current thesis, PAD formulations do not provide any substantial benefit over the distinctions identified in the circumplex model and so they are not explored here in greater depth.

Overall, the distinctions of the circumplex model map better onto the theoretical models of arousal and affect utilised in most cognitive ritual theories than do those of basic

emotion theories. Furthermore, there is robust and converging empirical evidence for there being a two factor dimensionality to affect and an association between emotional states and differential physiological activation. Consequently, for this thesis the circumplex model of affect is applied and related measurement tools are used to collect data based on its theoretical assumptions.

Turning then to examine the role of affect in group bonding, there has been growing interest recently on the function of shared affect in groups (Kelly, Iannone, & McCarty, 2014; Spoor & Kelly, 2004). In particular, shared affect in groups is argued to offer two coordination functions that have been crucial to human evolution: “communicating information among group members and creating social bonds that maintain the group” (Spoor & Kelly 2004: 408). A number of important mechanisms that enable affect to be readily transmitted through a group have also been identified including emotional contagion (Hatfield, Cacioppo, & Rapson, 1994) and interaction synchrony (Bernieri & Rosenthal, 1991; Chartrand & Bargh, 1999; Lakin, Jefferis, Cheng, & Chartrand, 2003). Collectively this research suggests that humans are highly sensitive to the affective cues of other group members and quickly respond to mirror the affective states of other members. A potential biological basis for such effects has also recently been uncovered through neuropsychological work on so-called *mirror neurons*; neurons that fire both when an action is performed by oneself and when observing someone else perform the same action (Gallese & Goldman, 1998; Iacoboni, 2009; Singer, 2006).

This work reflects a general renewed interest in research on evolution and the social role of emotions (De Waal, 1996; Robin I. M. Dunbar, 2003; Lieberman, 2007, 2013; Preston & de

Waal, 2002; Tooby & Cosmides, 1990). All of this research builds from the widely accepted foundation that different emotions serve different functions: a straightforward interpretation of which would be that negative affect is used to motivate individuals to avoid, or take action against, undesirable situations, whereas positive affect is used to reinforce desirable, fitness enhancing behaviours (Carver & Scheier, 1990).

In broad terms, this proves to be correct with evidence showing that positive emotions serve to facilitate social bonding (Shiota, Campos, Keltner, & Hertenstein, 2004), foster motivation (Isen & Means, 1983; Isen & Reeve, 2006), and increase success across multiple life domains, including marriage, friendship, income, work performance, and health (Lyubomirsky, King, & Diener, 2005). Conversely, negative emotions can be debilitating (Rugulies, 2002) and promote self destructive behavior (Stein, Apter, Ratzoni, Har-Even, & Avidan, 1998). However, negative emotions can be motivational, and act as a “common neural alarm system” for addressing physical and social pain (Eisenberger & Lieberman, 2004).

Due to such varied effects, there is ongoing debate about whether negative affect should be regarded as or orthogonal to, or independent from, positive affect (Green, Goldman, & Salovey, 1993; J. A. Russell & Carroll, 1999; Schimmack, 2008). For both positive and negative affect, the precise hierarchical structure of constituent elements also remains under debate (Bagozzi, 1993; Crawford & Henry, 2004; Tellegen et al., 1999; Watson & Clark, 1994). However, regardless of which theoretical model is endorsed, the majority of research on negative affect has focused on its damaging impact on individuals and within a group context its role in fostering “hostility and collective action towards enemies and

predators” (Spoor & Kelly 2004: 405). Spoor & Kelly in an influential review of the literature on affect within groups posited that the role of negative affect in a group setting “would seem to be most related to the communication function”, whereas positive affect “would seem to be most related to the bonding function” (2004: 406).

Such an interpretation is at odds with Modes theory and contradicts cognitive dissonance research but it may also reflect an outdated view, as a more recent summary by Kelly et al. (2014: 180) notes that there is growing recognition that shared negative affect might also serve as a means to “solidify group bonds” and not just through enhancing intergroup hostility. For example, in a study examining interactive and non-interactive affective sharing in work groups, Klep, Wisse, & Van Der Flier (2011) found that those in the interactive negative affective condition reported more belongingness than those in the positive mood condition. Similarly, Grawutch, Munz, & Kramer reported that when working on a creative production task group, members who experienced negative affect displayed a stronger relational focus. Furthermore, there is a robust association documented in the psychology literature between fear and social affiliative tendencies (Darley & Aronson, 1966; Dunn & Hoegg, 2014; Kirkpatrick & Shaver, 1988; N. Miller, 1966; Teichman, 1973; Werth, Steidle, & Hanke, 2012; Wirth & Schultheiss, 2006).

These results better accord with the predictions of Modes theory and they are indicative of a wider re-assessment of the role of negative affect in group bonding, which is discussed in greater detail in the following section.

1.9 Dysphoria & Social Bonding

As already discussed participation in collective dysphoric events, including ritual performances, has been theorised to provide an array of benefits in relation to forging bonds between performers and within a community. The paradox that this represents has recently begun to be addressed in a number of studies that have sought to examine “the positive consequences of pain” (Bastian, Jetten, & Ferris, 2014; Bastian, Jetten, Hornsey, et al., 2014; Leknes & Bastian, 2014; Olivola & Shafir, 2013; Xygalatas et al., 2013b). In a recent review paper covering material from a range of academic fields, Bastian et al. (2014) outline nine positive effects of pain grouped into three main categories: 1) Pain facilitates pleasure; 2) Pain enables self-regulation and enhancement; and 3) Pain promotes affiliation (Table 1.3). The third category identified is of particular relevance to this thesis and the connections posited between imagistic ritual practices and identity fusion, as the evidence reviewed supports the view that experiencing pain promotes a relational focus and a sense of solidarity within a group. Moreover, the evidence for the second category of benefits provides support for both costly signalling theories (Bulbulia, 2004a; Henrich, 2009) and the cognitive resource depletion model (Schjødt et al., 2013) as it illustrates how enduring pain can act as a strong signal of virtue (D. B. Morris, 1991; Schneider, 2009) and promote a focus on low-level representations serving as a vehicle to make individuals “open to novel... representations” (Bastian, Jetten, & Ferris, 2014: 263).

The suggestion of an affiliative role for pain may seem counterintuitive given that pain typically serves as a reliable threat-signal and should encourage individuals to withdraw. However, the fundamental desire to escape from pain appears to be part of what motivates

individuals to increase their relational focus, as they search for sources of social support to either aid their escape or alternatively just to provide comfort. In support of this, perceived social support was found to predict pain adjustment among patients suffering from chronic pain (López-Martínez, Esteve-Zarazaga, & Ramírez-Maestre, 2008) and to correlate with lower self-reported labour pain during childbirth (Cogan & Spinnato, 1986; Lidderdale & Walsh, 1998). Moreover, laboratory based studies that ask participants to endure various types of experimentally induced pain (for an indepth review see Krahé, Springer, Weinman, & Fotopoulou, 2013) have found that participants report less pain when they are reminded of partners (Master et al., 2009; Younger, Aron, Parke, Chatterjee, & Mackey, 2010), salient group identities (Jones & Jetten, 2010; Michael J. Platow et al., 2007), or when they are provided with social support, regardless of the identity and level of intentionality of the provider (J. L. Brown, Sheffield, Leary, & Robinson, 2003). One study also found that religious images were capable of reducing the perception of pain for religious believers but not for atheists and agnostics (Wiech et al., 2008). Finally, a study of athletic rowers found greater pain endurance immediately after group training session as compared with performance after individual training (Cohen et al., 2010).

All of this is indicative of the fact that pain is not simply an individual experience but can be significantly modulated by the social environment. Yet, even though there is strong ethnographic evidence for the role of pain, or broader dysphoria, in increasing social cohesion there are currently very few experimental studies that have directly addressed this question. A recent exception is a study by Bastian, Jetten, & Ferris (2014) which reports the result of three innovative experiments that address whether painful experiences can

promote cooperation with social groups. The findings of this study will now be examined in depth as first, they provide a useful illustration of how dysphoric factors can be manipulated in collective settings under laboratory conditions and second, the findings were explicitly presented as offering evidence of the ‘social glue’ theorised recently by Modes theory (Whitehouse & Lanman, 2014; Whitehouse, 2013c).

Table 1-3: The benefits of pain and associated processes

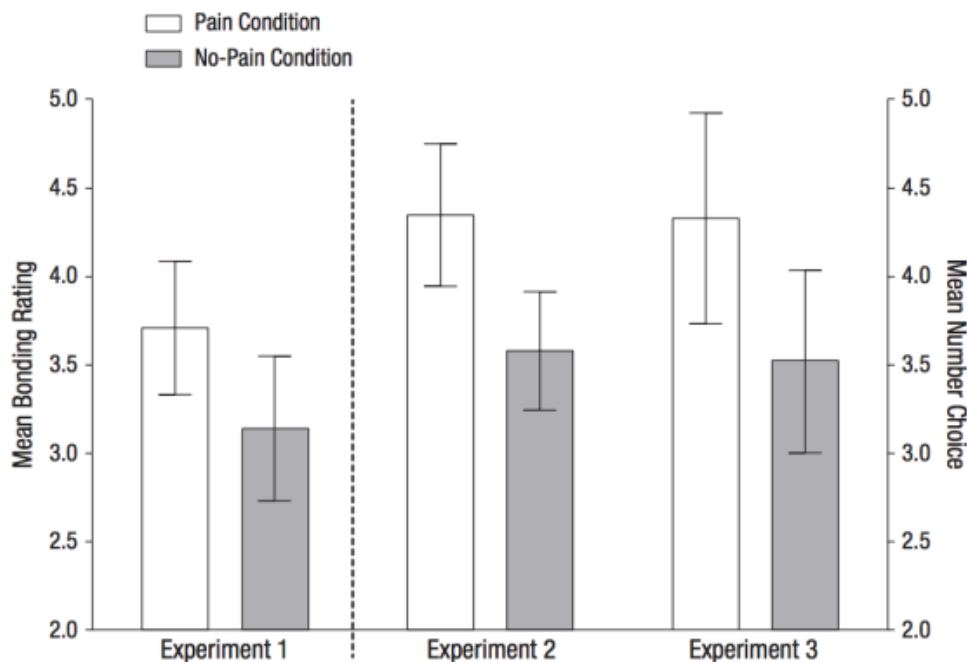
Benefits	Process
1. Pain facilitates pleasure.	
i) Pain enhances subsequent pleasure.	Pain provides a contrast for pleasure and this increases the relative pleasantness of subsequent experiences.
ii) Pain heightens sensory sensitivity.	Pain heightens arousal and constrains attention on sensory experience, thereby increasing sensory receptivity.
iii) Pain facilitates pleasure seeking.	Pain provides a justification for indulgence of pleasures that might otherwise arouse a sense of guilt.
2. Pain enables self-regulation and enhancement.	
i) Pain increases cognitive-affective control.	Pain captures attention and brings cognitive resources on-line for effective problem solving in response to the threat of pain.
ii) Pain enables identity management	Pain promotes a physical experience of the self, thereby reducing high-level self-awareness and enabling identity change.
iii) Pain demonstrates virtue	Pain may be interpreted as providing a symbolic test of a range of personal virtues.
3. Pain promotes affiliation	
i) Pain arouses empathy in others	The expression of pain increases empathy and arouses care and concern in others.
ii) Pain increases relational focus	People seek social support in response to pain. Pain therefore provides a novel source of social connection with others.
iii) Pain increases solidarity	Pain may be used to increase the value of relational bonds with others, and shared pain may increase interpersonal bonding.

Source: (Bastian, Jetten, Hornsey, et al., 2014)

Each of the three experiments in the study involved comparing aspects of ‘group bonding’ after the performance of similar tasks that differed only in whether they involved pain. In the first two experiments the two conditions [pain and no pain] were created

through the sequential performance of two tasks, which involved individuals first submerging their hands in water that was either very cold ($<3^{\circ}$) [Pain Condition] or at room temperature ($>24^{\circ}$) [No Pain Condition]; ostensibly to complete a sorting task and second, performing a wall squat “with back straight and knees bent at 90° ” for as long as possible [Pain Condition] or an attempt to balance on one leg for 60s (with the ability to switch legs if tired and/or use balance aids) [No pain condition].

Figure 1-11: Differences in bonding and pro-group behaviours over three experiments



Source: (Bastian, Jetten, & Ferris, 2014: 2081)

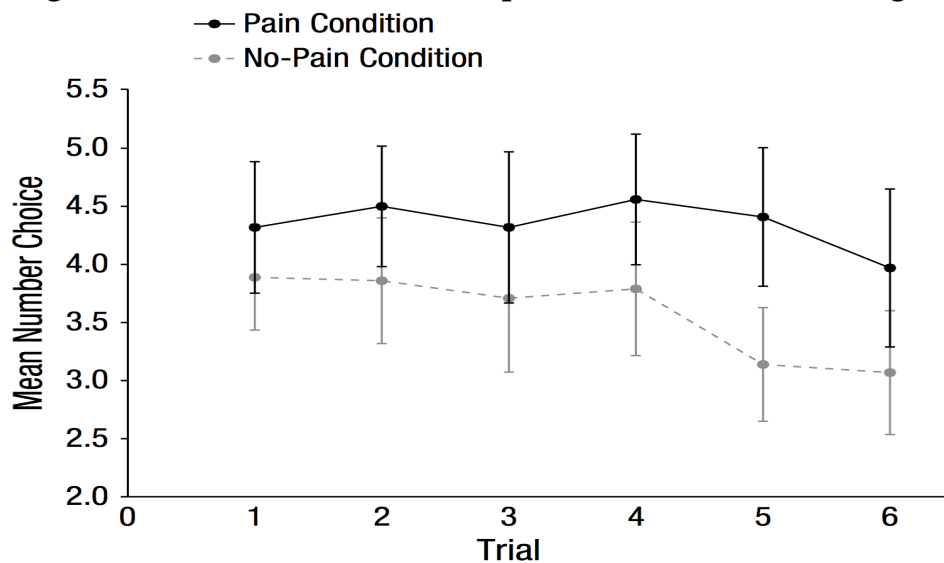
These tasks were each performed individually but in the presence of a small group (2-5 people), who also performed the same tasks. Manipulation checks confirmed that those who experienced the cold water and painful wall squat endurance task regarded their experiences as more painful and more unpleasant than those who performed the warm water and balancing tasks. The observed outcome was that those who endured the painful version of the tasks also reported feeling more bonded with the other participants, as

measured by how far they agreed with statements like: “I feel connected to the other participants” and “I feel like there is a unity between the participants in this study”. A medium-sized effect was observed ($d=.054$) and this remained when controlling for other possible confounding factors such as age, gender, and group size (Fig. 1.11).

For the second experiment, instead of a self-reported measure of bonding a ‘weakest link’ economic game (Hirshleifer, 1983) was used to provide a behavioural measure of the consequences of ‘bonding’. The rules of the game specified that each participant should select a number between 1-7, with the biggest possible payoff (\$7.80) for all participants available if everyone in the group selected 7. However, if a participant selects a high number while some other member of the group selects a lower number, they risk receiving a reduced amount. For instance, if one participant selects 7 and another member of the group selects 1, regardless of the choices of any other members, the participant who selected 7 would only receive (\$0.60) while the defecting member receives (\$4.80). This structure means that the safest strategy is to always select 1 as then your reward amount is guaranteed to be at least \$4.80 as no other group members can select a lower number. However, if you were able to trust the other members within the group the more rewarding strategy is to select the highest option 7, as this offers the largest individual and collective reward when no other group members defect by selecting a lower value.

Using this game, played over six rounds a significantly higher rate of cooperation was found in the pain condition (Fig. 1.12). Controlling for age and gender did not impact this result but when controlling for group size the p-value of the observed difference decreased from $p < .001$ to $p = .048$ which reduces confidence in the validity of the finding.

Figure 1-12: Ex.2: Differential cooperation in a weakest link game



Source: (Bastian, Jetten, & Ferris, 2014: 2082)

Fortunately, there was a third experiment that employed essentially the same structure, including the weakest link economic game as an outcome, providing an independent test of the finding from the second experiment. The results obtained replicated the same elevated cooperation levels in the pain condition (Fig. 1.11) and, importantly, there was no average group size confound between conditions in this sample. The third experiment also employed an entirely novel means of manipulating pain, that involved the consumption of hot chilli peppers in the pain condition and hard candy in the control condition under the cover story of a consumer preference test. This new framing also meant that there was less chance that the results were due to factors associated with competition or perceptions of 'success' from the sorting and endurance tasks.

In conclusion, all three experiments reported demonstrated a consistent pattern of pain increasing bonding and pro-group behaviour (Fig. 1.11) with the effects not restricted to a single variety of painful stimuli. One limitation of the findings however is that all of the

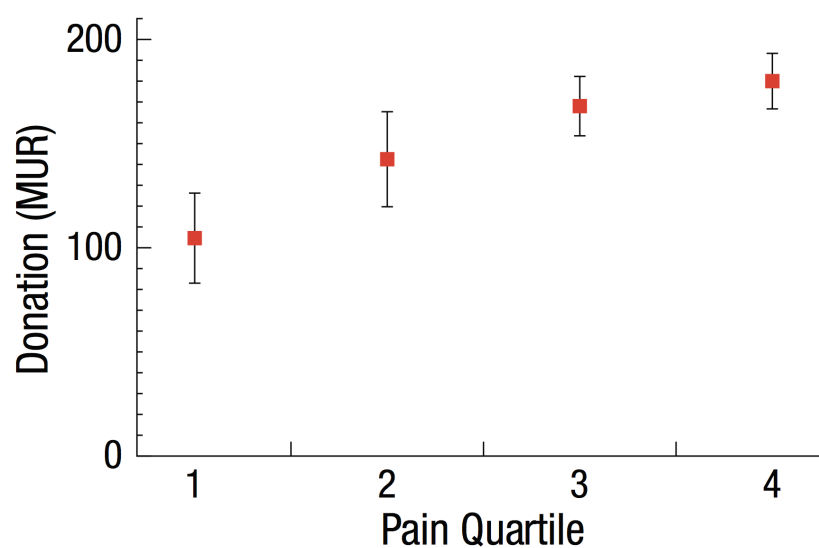
experimental treatments manipulated varieties of physical pain and thus it is unclear whether similar effects observed could also be anticipated with shared psychological pain. The authors note that since the painful tasks were framed as individual rather than group tasks “the enhanced bonding and cooperation... emerged from the experience of the pain rather than... pain for the group” (Bastian, Jetten, & Ferris, 2014: 84). This is important as it makes an explanation based on an increased evaluation of the group due to motivations to reduce dissonance unlikely (Aronson & Mills, 1959; Festinger, 1962; Gerard & Mathewson, 1966). Additionally, the artificial nature of the group and lack of salient pre-established social identities mean that costly signalling accounts are also a poor fit for the observed results. Finally, the fact that participants in the control condition participated in functionally similar tasks but did not show the same elevation in bonding or cooperation suggests that sharing a common fate with group members is also insufficient to produce the effect.

The studies’ design did not provide a means to detect a mechanism for the effects observed but the authors suggested that a plausible pathway was that “pain promotes cooperation because of its well-documented capacity to capture attention and focus awareness on the immediate painful event” which, in a shared painful experience, “not only makes the events more salient but also enhance the salience of the other people who shared in these events” (Bastian, Jetten, & Ferris, 2014: 2084). This account aligns with Modes theory model of imagistic rituals, particularly in relation to the proposals that dysphoric rituals to prompt deep reflection (Richert et al., 2005), generate richer internal

representations, and promote stronger relational bonds with those who share such self-defining moments (Whitehouse & Lanman, 2014).

This overlap was also explicitly recognised by Bastian et al. (2014: 2084) who commented that their interpretation “aligns with the accounts of Whitehouse and his colleagues” but at present this interpretation remains speculative. Further research is required to establish that such artificial painful experiences could have any significant impact on self-concepts. However, an interesting real world correlation with these findings comes from Xygalatas et al.'s (2013b) research on the dramatic Hindu Thaipusam festival in Mauritius which involves devotees piercing their skin with multiple “needles, hooks and skewers”. This study is discussed in detail in the next review section but a finding particularly relevant to the experiments just discussed, was that costly donations to a relevant Hindu temple were significantly correlated with the perceived pain of the task for both participants in the ritual and observers.

Figure 1-13: Donations increase with higher perceptions of pain amongst devotees



Source: (Xygalatas et al., 2013b: 2)

The results from the study by Bastian et al. and the field data from Mauritius collectively offer preliminary evidence that sharing pain collectively can increase pro-group behaviour and enhance the sense of solidarity between group members, even when the pain is not personally experienced.

If these findings prove to be correct then they add a new facet to the existing literature on cognitive dissonance (Festinger, 1962) and the role it plays in group initiations (Aronson & Mills, 1959; Hautaluoma & Spungin, 1974; Kamau, 2013; Lodewijkx, 2005). The classic illustration of the dissonance effect is from Aronson & Mills (1959) study, which varied the severity of the initiation required to gain access to a discussion group on the psychology of sex. This study had three conditions with different degrees of severity for the initiation test required to gain entry to the group.

Significantly, the dysphoric element of this initiation test was not related to physical pain but rather social discomfort and embarrassment. In the severe condition, participants were required to read 12 obscene words aloud to the experimenters and two explicit sexual descriptions from a novel. Whereas in the mild condition they read 5 words related to sex but which were not as obscene. Finally, in the control group there was no initiation test. In each condition following the 'initiation' the participants were then allowed to listen to an on-going group discussion held in another room, but in reality this was the same tape-recorded conversation across all conditions, designed to be "one of the most worthless and uninteresting discussions imaginable" (Aronson & Mills, 1959: 179). Following this, participants were asked to rate the discussion and the other participants on a variety of scales. The results revealed that those who had undergone the more severe initiations

reported a much higher level of liking for both the content of the discussion and the other group participants (Fig. 1.14) than in both the mild and control conditions. Aronson & Mills suggested that the explanation for this outcome was that participants in the severe condition had sought to reduce the sense of dissonance created through completing a painful task for a dull group by enhancing their evaluation of the group.

Aronson & Mills experiment manipulated the level of embarrassment in their initiation conditions but described it as representing 'pain' but follow up studies demonstrated that the findings were supported when initiations involved discomfort (Keating et al., 2005, Study 3), painful electric shocks (Gerard & Mathewson, 1966), drawing blood, vomiting and injury (Winslow, 1999) or distress (Hautaluoma & Spungin, 1974). Alternatively, employing a frustrating math task Kamau (2013) failed to find any increase in either group identification or liking²³. The replication by Gerard & Mathewson (1966) involving painful electric shocks is of particular significance, as it established that the effects of increased liking were only observed when they were presented as a necessary pre-requisite of joining the group. This suggests that the findings cannot be solely attributed to the effects of individual painful experiences and instead relate to the specific social elements of the experience.

A recent set of studies that examine how enduring pain relates to perceptions of morality and justice also offer evidence for how social factors can strongly influence the interpretation of painful experiences. Bastian, Jetten, & Fasoli (2011) demonstrated that

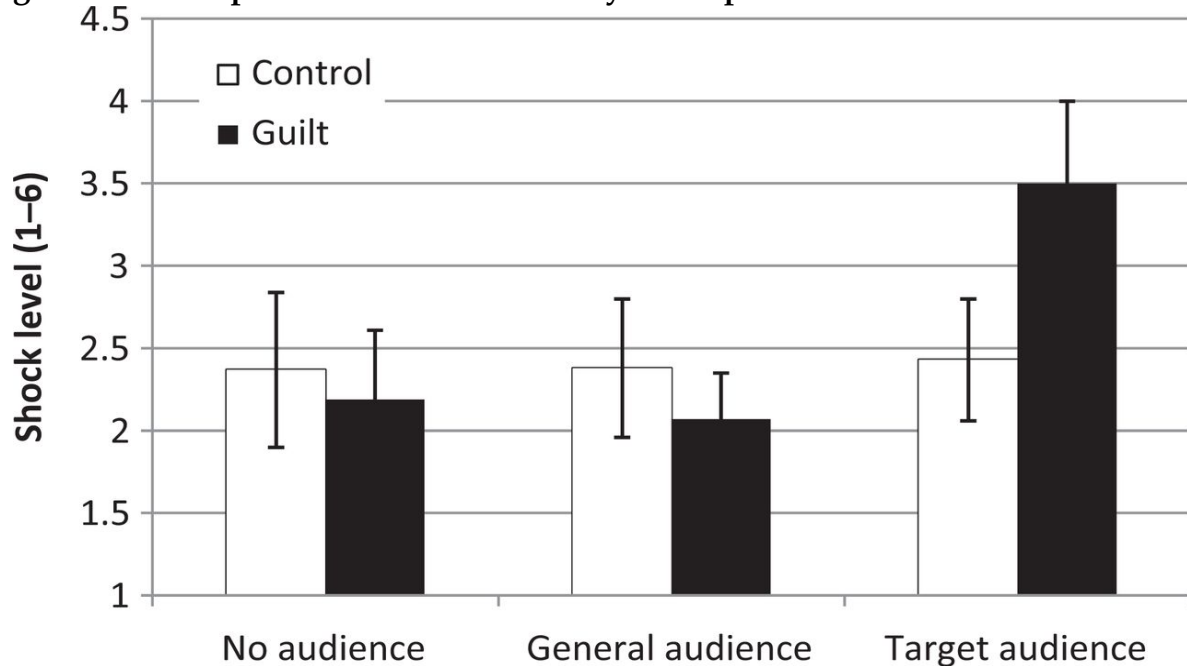
²³ This finding was likely confounded by the inclusion of a reward element that may have helped to reduce dissonance. The study also had no comparative control condition and did not collect any manipulation checks for negative affect, meaning that it is difficult to determine whether there were any significant differences created between the mild and severe frustration conditions.

when participants were reminded of their past immoral deeds they subjected themselves to more pain by keeping their hands submerged in an ice bath (Walsh, Schoenfeld, Ramamurthy, & Hoffman, 1989) for longer periods in a purported physical acuity task. Additionally, those who experienced pain were found to have reduced feelings of guilt, as compared to those who performed the same non-painful task with warm water and a control group. While these results support a link between pain and morality they also relate to sociality as the immoral deed that participants were asked to recall was “a time when they rejected or socially excluded another person” (Bastian et al., 2011).

In contrast to this another study by Van Bunderen & Bastian (2014) found that participants who were led to believe they had performed unfairly²⁴ in a two player cooperative performance task displayed *reduced* “justice-restoring behavior” when they had previously completed the same painful activity (immersing their hand in cold water) as compared with those who had completed the non-painful version of the same activity. This result might initially appear contradictory of the finding from Bastian et al. (2011) but the pain endured seems to again have been interpreted as a cleansing experience. Consequently, by reducing the participant’s sense of moral debt resulted in a decrease in the amount of voluntary compensatory action undertaken. Although enduring pain was not framed in the experiment as being related to the other participant or choices in the cooperative task and was correlated with less prosocial behaviour it does seem to have been interpreted as a means of addressing a social debt.

²⁴ This effect was however only found amongst participants who were sensitive to personal injustice.

Figure 1-14: Self-punishment increases only in the presence of those treated unfairly.



Note: Error bars represent standard errors.

Source: (Nelissen, 2011: 142)

This interpretation is supported by the findings of another study by Nelissen (2011), which used a similar two-player game to induce feelings of guilt from an unfair performance and then subsequently gave participants the opportunity to select the severity of electric shock that themselves would receive for failing to provide correct answers in an unrelated quiz. The key finding from this study was that participants selected a stronger setting for their electric shock punishment when they i) felt guilty for performing unfairly and ii) believed their partner from the previous task would be present in the audience (Fig 1.15). This result demonstrates clearly that enduring pain can serve as a sign of interpersonal remorse and reduce a moral debt, even when performed as part of an unrelated task. The results also emphasise the significance of surrounding social factors, since participants were only motivated to endure additional pain when they believed that those they had treated unfairly would be present to observe their actions.

The ability of dysphoric experiences to serve as a social signal of group commitment has already been discussed in relation to costly signalling and Modes theory but these theories have primarily addressed such effects in religious contexts. Independent from these studies there has also been substantial work conducted examining the motivations and consequences of hazing practices in University fraternities (Campo, Poulos, & Sipple, 2005; Cimino, 2011, 2013; Johnson, 2011; Keating et al., 2005; Ramey, 1981; Walker, 1967), sports clubs (E. Anderson, McCormack, & Lee, 2011; Finkel, 2002; Groves, Griggs, & Leflay, 2011; McDonald & Sylvester, 2013; Van Raalte et al., 2007), and military organisations (de Albuquerque & Paes - Machado, 2004; Østvik & Rudmin, 2001; Pershing, 2006; Winslow, 1999). Although some of this research is focused on regulation and normative interventions (Holmes, 2013; Nuwer, 1990, 2001, 2004; Waldron, 2012) and subject specific concerns, a broad evolutionary informed approach called *automatic accrual theory* has been developed by Aldo Cimino (Cimino & Delton, 2010; Cimino, 2011, 2013; Delton & Cimino, 2010).

Cimino (2013: 446) defines hazing as “the abuse of new or prospective group members”²⁵ and notes that although the content of hazing is “profoundly variable” there are four consistent features of the behaviour (Cimino, 2013: 447), specifically hazing is: 1) *Temporary*: Hazing has a restricted time course and is rarely repeated; 2) *Unidirectional*: Hazing is solely directed at newcomers by veterans; 3) *Coercive*: Hazing is inflicted and is often presented as an inescapable social obligation; 4) *Coalitional*: Hazing arises in long-term cooperative alliances.

²⁵ Cimino (2013: 446) also provides a longer definition for hazing as “the generation of inductions costs (i.e. elements of the experiences necessary to be acknowledged as a ‘legitimate’ group member) that appear unattributable to group relevant assessments, preparations or chance”.

Cimino (2013: 447) argues that these specific features cannot be accounted for via the common explanations and in particular cognitive dissonance accounts are problematic: “if effort justification can increase group liking, why not just continue hazing? Further... [why do] veterans [not] have to justify their effort by being abused by newcomers? ... why is it [only] concentrated within coalitional groups?”. Hazing events are also typically coercive experiences but removing the perception of choice is known to diminish feelings of dissonance (Harmon-Jones & Harmon-Jones, 2007).

As an alternative to dissonance accounts Cimino proposes *automatic accrual theory* which argues that a core reason why hazing occurs is because “the human mind is equipped with psychological mechanisms that motivate the strategic devaluation of coalition newcomers” (2013: 447). This mechanism is suggested to have evolved due to the key adaptive problem faced over evolutionary history of accurately assessing the intentions and quality of new group members. Specifically, Cimino highlights that enduring coalitions with multiple, overlapping membership generations are often able, over time, to amass substantial group benefits which includes common goods, property, and status. The issue then is how these kind of groups can prevent exploitation of their resources from non-contributing free riders. Of particular concern is that while access to certain resources can be restricted by achievement or seniority, there are inevitably certain benefits that are accrued *automatically* from group membership, including status, protection, and access to common property. This means that coalitions are vulnerable to exploitation from newcomers who may only temporarily associate with the coalition to extract automatic benefits with no intention of contributing to the costs of the group or alternatively associate over the long term but

provide little or no additional value to the group. The strategic devaluation of newcomers resolves these concerns by dramatically increasing the costs associated with group entry and temporarily enforcing compliance and conformity to emphasise the dominant position of veteran members.

Evidence in support of this theory has been provided by vignette experiments that ask participants to imagine themselves as members of strongly cooperative (the Ice Walkers) or weakly cooperative (the Bug Watchers) groups and to decide on the initiation practices for new members. In both a university sample (Cimino, 2011) and a representative sample of the public from the United States (Cimino, 2013) participants that imagined themselves as members of strongly cooperative, enduring coalitions desired more severe initiation practices than those who imagined themselves as part of less cooperative groups. Moreover, the greater the level of perceived automatic benefits from group membership the more initiation practices were desired; with males also desiring more *severe* initiations (Cimino, 2013: 450). This correlation was not found for non-automatic earned benefits. Even though these two studies used hypothetical groups, there have been other studies conducted with real groups that reveal the same patterns. For instance, Walker (1967) surveyed the prestige (an automatic benefit) of 29 fraternities at the University of Washington and found a positive association with hazing severity. Similarly, Ramey (1981) examined 31 different fraternity chapters and found a positive association between the perceived toughness of their induction process and their prestige. Finally, from ethnographic data, Young's (1965) study of male initiation rites found that societies with

more powerful coalitions tended to have more 'dramatic' initiation practices and a similar pattern was reported for mystery cults in Melanesia (Allen, 1967; Strathern, 1970).

When the experimental and observational data is combined the evidence in support of automatic accrual theory is preliminary but compelling however it is also evident that in order to serve as a convincing evolutionary model, there is a need for stronger cross-cultural evidence that the trends observed with US samples are more broadly applicable. It is also important to note that this theory cannot account for dysphoric practices that are not performed by newcomers as part of initiations, which Atkinson & Whitehouse (2011) revealed to account for a substantial proportion of ritual practices in their worldwide sample.

The selection of studies reviewed above represent the major cognitive theories addressing the bonding effects of dysphoric practices. Pain is the dysphoric experience that has received the most attention and is also potentially the most relevant for dysphoric ritual practices. However, the evidence also suggests that other types of dysphoria, such as embarrassment (Aronson & Mills, 1959; Gerard & Mathewson, 1966) and fear can produce significant affiliative effects (Graham, Huang, Clark, & Helgeson, 2008). Although there is certainly a need to differentiate between different sources of dysphoria and their specific effects (Czachesz, 2010), it is also important to recognise that in practice there will often be substantial overlap in the components of negative affect, anxiety often accompanies pain for example, that make it appropriate to consider responses to the general category of dysphoria. In the following section, this inclusive stance is adopted in order to review

experimental studies that have examined dysphoric rituals and their impacts on bonding processes.

1.10 Dysphoric Ritual Studies

In anthropology and religious studies, the most influential research on ritual has traditionally been based on first-hand ethnographic observation of ritual behaviour. Detailed ethnographic descriptions from such studies have also been utilised in coding studies by CSR researchers (Atkinson & Whitehouse, 2010; Sosis et al., 2007). However, as the CSR field has developed interest has grown in exploring rituals more precisely through the use of experimental studies, both in controlled lab environments and at actual ritual events. At present, this remains a new area of research and as a result there are only a small number of existing studies to draw from. This section reviews the existing experimental literature focusing particularly on studies addressing dysphoric rituals.

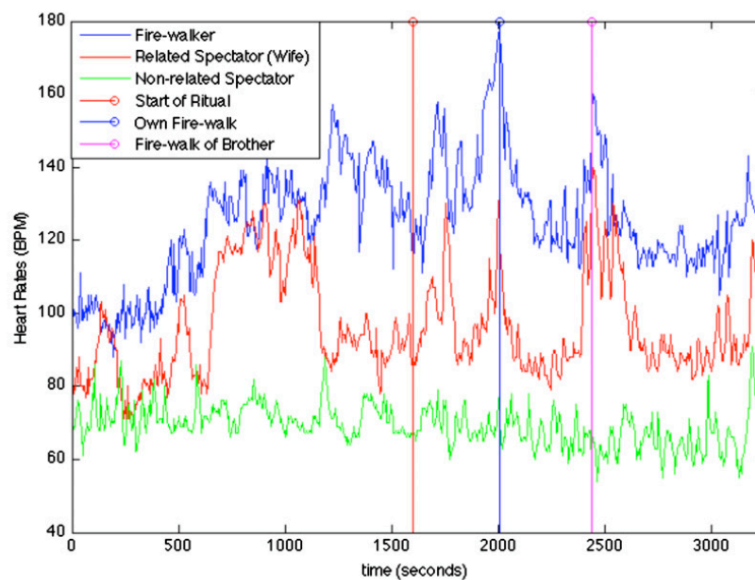
Experimental studies of rituals can be separated into two categories: field experiments and artificial laboratory experiments. Field experiments are opportunistic in that they rely on the observation and measurement of naturally occurring rituals and attempt to make use of existing divisions (for example, between ritual participants and observers) in order to establish comparative categories and test hypotheses. Traditionally such studies would have relied primarily on the researcher's observational skills but more recently the use of new technologies and field experiment methods have enabled researchers to gather more precise measurements (Cohen, 2007; Fischer et al., 2014; Fischer, Callander, Reddish, & Bulbulia, 2013; Purzycki & Arakchaa, 2013; Richert et al., 2005; Whitehouse et al., 2014; Xygalatas et al., 2013b).

The research of Dimitris Xygalatas serves as a paradigmatic example of field-based research that employs quantitative methods. His initial ethnographic research focused on fire-walking rituals, known as the *Anastenaria*, practiced in rural Greece (Xygalatas, 2012) and argued for the importance of addressing psychological and neurobiological factors, in addition to the social and historical elements, to understand the appeal of the practice and the effects of participation. Building on this experience Xygalatas and a team of researchers conducted an experimental study (Konvalinka et al., 2011; Xygalatas et al., 2011) that set out to quantify the physiological effect of participation and observation of a dysphoric firewalking event held in the small Spanish municipality of San Pedro Manrique. The primary goal of the study was to address whether such a highly arousing ritual event was capable of generating shared emotional states, what Durkheim (1965) referred to as *collective effervescence*, despite the ritual lacking any synchronous movement (firewalkers crossed the coals one at a time). To attempt to quantify collective effervescence concretely the researchers took heart rate measurements by attaching sensors to both firewalkers and observers during the ritual event.

The results revealed clear synchronous arousal patterns between firewalkers and observing family members/friends in the audience but not with unrelated members of the audience (see Fig. 1.16). This was a significant finding firstly, because it offered evidence that the synchronous performance of actions in collective rituals was not the only mechanism to generate a shared affective response and secondly, it implied that to experience synchronous arousal attendance at a ritual is not sufficient and instead “the observers must share membership in the group and have a relationship with at least one of

the participants” (Konvalinka et al., 2011: 8518). A potential issue for this interpretation is that there was also greater synchrony in the baseline heart rates of related individuals collected a number of hours before the ritual began. This could be, as the authors suggest, due to the fact that “the anticipation of the ritual carries structure as well” but it could also indicate that the synchrony observed was due to other factors not related to the ritual (Konvalinka et al. 2011: 8515).

Figure 1-15: Illustration of heart rate patterns during a firewalking ritual



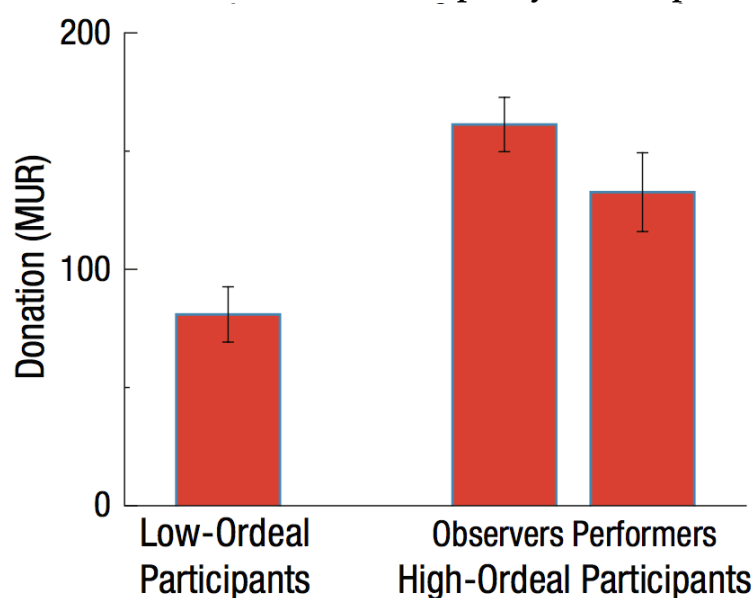
Source: (Konvalinka et al. 2011: 8515).

Another recent study by Xygalatas et al. (2013b) sought to address whether shared empathic arousal at dysphoric ritual events translated into prosocial behaviour and a more inclusive self-identity (Tseung-Wong & Verkuyten, 2010). To do so Xygalatas and a team of researchers attended the Hindu *Thaipusam* festival held in Mauritius, during which devotees engage in dysphoric practices “involving body piercing with multiple needles and skewers, carrying heavy bamboo structures, and dragging carts attached by hooks to the skin for over 4 hrs. before climbing a mountain barefooted to reach the temple” (Xygalatas

et al., 2013b: 1602). This study is described in depth as the methods and findings pertain directly to the field studies undertaken in this thesis.

During the Thaipusam festival alongside the high arousal piercing rituals there are also low arousal rituals involving collective prayer. The researchers used this natural variance to collect responses that could serve as a low arousal comparison group. Importantly, they noted that “all participants [eventually] took part in both rituals” and consequently, although they only collected responses from each participant at one of the events, they avoided the problem of self-selection bias. Prosociality was operationalised on two measures: i) a behavioural measure of anonymous charitable donations to the temple and ii) a self-report measure of social identification measured on an 18-point scale with the more “parochial ethnic-religious Hindu identity” on one pole and “the superordinate inclusive national Mauritian identity” at the other (Xygalatas et al., 2013b: 1603).

Figure 1-16: Mean donations to temple by festival participants

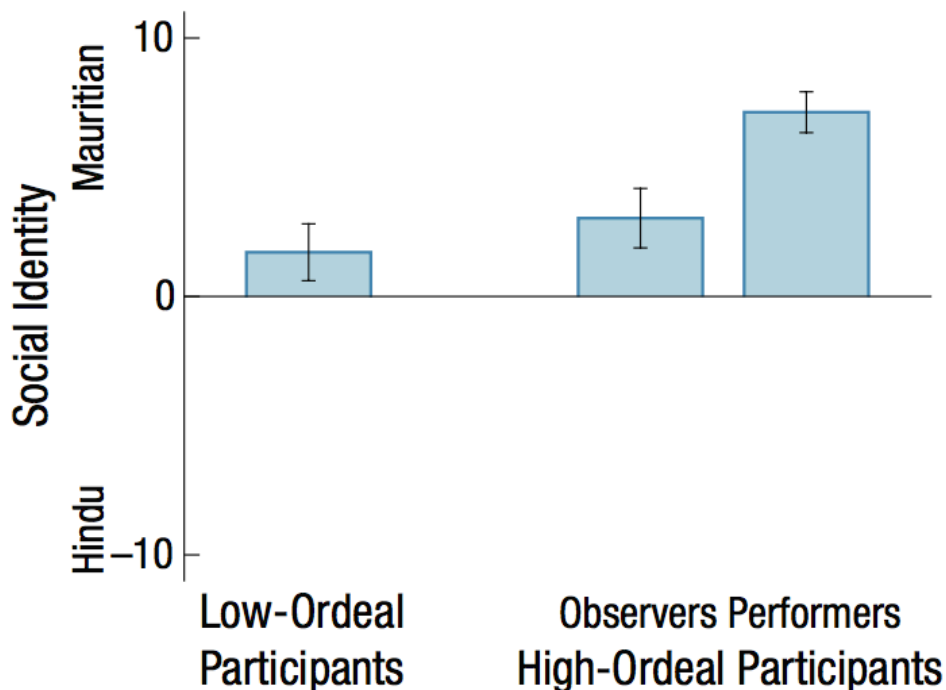


Note: Error bars represent standard deviation.

Source: (Xygalatas et al., 2013b: 1082)

The finding that perceived pain predicted donations has been discussed in the preceding section but in regards to the two key measures of prosociality: first, it was observed that both high ordeal observers and performers donated significantly more than low ordeal performers (there were no non-participants in the low ordeal ritual) (Fig. 1.17) and second, although all participants endorsed the more inclusive Mauritian social identity, this effect was found to be strongest for high ordeal performers (Fig. 1.18).

Figure 1-17: Endorsement of exclusive (Hindu) vs. inclusive (Mauritian) identities



Note: Error bars represent standard deviations.

Source: (Xygalatas, 2013: 1062)

These results collectively provide support for the general association proposed between prosociality and religion (Norenzayan et al., 2013; Norenzayan & Shariff, 2008; Shariff et al., 2015) and also suggest that prosociality can be elevated by both observing and performing dysphoric ritual practices. What is more, although there was no significant difference

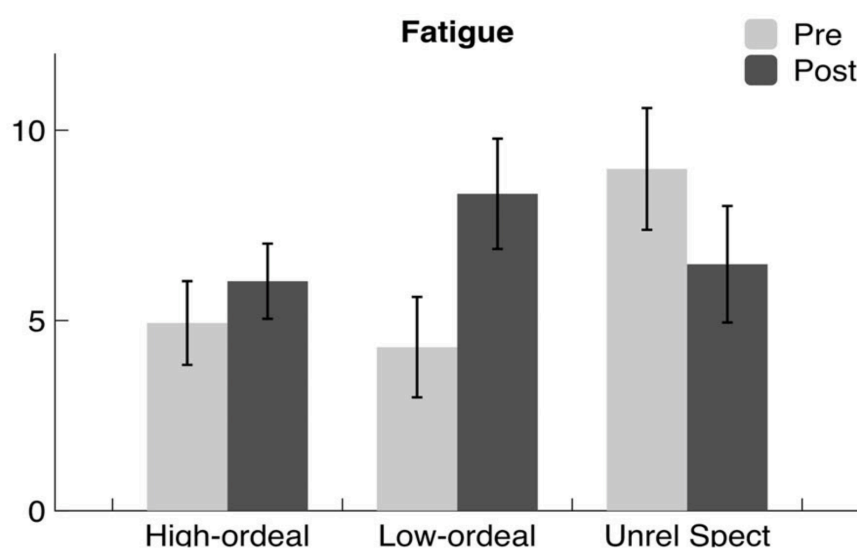
detected between the donations of high ordeal observers and participants ($p=.15$), there was a visible trend for greater donations amongst observers (Fig. 1.17).

This trend could be the result of random variation but it is difficult to assess due to the relatively small sample sizes of the two groups; which means the study may have been underpowered to detect a smaller effect. At present there is not enough evidence to resolve the issue of whether the trend observed was random variation, however if this difference was replicated in subsequent studies it would provide some support for costly signalling interpretations. In particular, greater prosocial activity may be elicited amongst observers who witness other group members performing a CRED and physically demonstrating their commitment to their ingroup.

However, another plausible interpretation could be that the endurance of pain acts serves as a moral cleanser reducing an individual's perception of moral debt and therefore enabling them to donate less to their a morally relevant cause or group. But this interpretation fails to address the relative elevation in prosocial behaviour observed in the performers in the high ordeal group as compared with those in the low ordeal group. This difference can also not be attributed to selection bias since all performers included in the study ultimately performed both the low and high ordeal rituals.

One implication of the elevated prosociality of observers is that the social bonding induced by publically shared pain (Bastian, Jetten, & Ferris, 2014) appears to extend to some to all those present at dysphoric rituals, assuming that they are members of the relevant in-group involved with the ritual performance.

Figure 1-18: Self-reported fatigue in a collective firewalking ritual



Note: Error bars represent standard deviation.

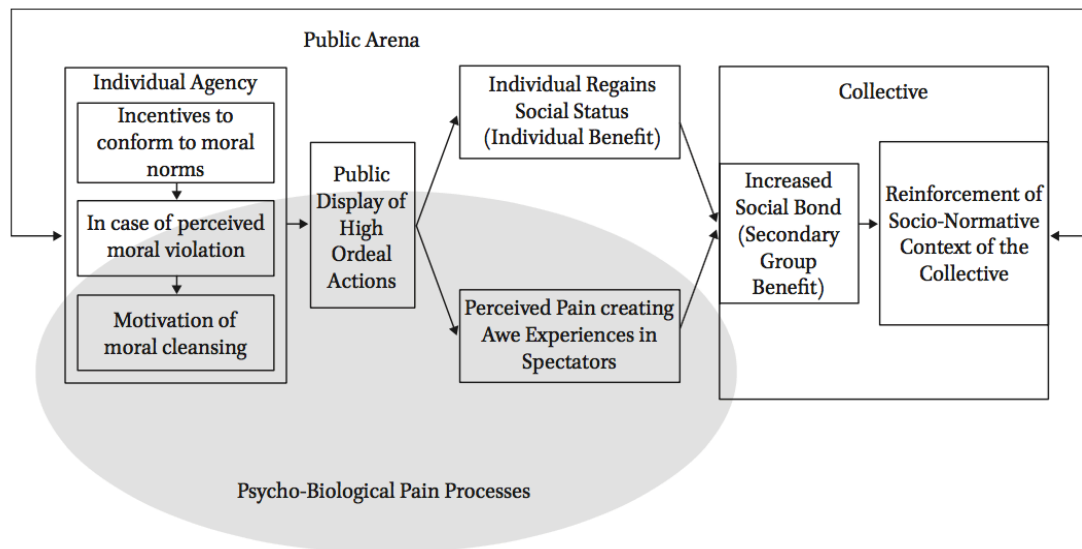
Source: (Fischer et al., 2014: 4)

This interpretation received additional support from the findings of a follow up study conducted at a collective fire-walking ritual performed in a Mauritian Hindu community (Fischer et al., 2014). This study adopted physiological measures similar to those used in the Spanish fire-walking study (Xygalatas et al., 2011) but retained comparisons between high-ordeal performers (fire walkers), low-ordeal participants (observers with familial bonds with performers), and unrelated spectators. Although physiological measures showed the anticipated elevation in heart rates for performers, the most significant findings were obtained from the self-reported measures which indicated that in the period after the firewalking performance, ritual performers reported greater levels of happiness than all observers, and conversely related observers (labelled low ordeal participants) reported higher levels of fatigue (Fig.1.19). This finding serves as an important reminder that dysphoric ritual events may not necessarily be associated with elevated levels of negative affect from performers and also indicates that observers with familial ties with performers

can experience negative affect and emotional exhaustion collectively through the experience (Fischer et al., 2014). This would explain why the prosocial effects of high-ordeal rituals are also be found with observers (Xygalatas et al., 2013b) even without a costly signalling interpretation.

Despite these findings it should also be noted that there are always other potential features of collective rituals that provide alternative and confounding pathways to prosocial behaviour. For instance, in another field study conducted with nine communities with distinct 'ritual types' (Fischer et al., 2015: 11), elevated prosociality was found in "rituals with synchronous body movements" and those with higher "levels of sacred attribution". Similar to the methods of Xygalatas et al. (2013b), prosociality was measured through self-reported attitudes towards other participants and decisions in a public goods game (Wiltermuth & Heath, 2009). This study serves as an important reminder that ritual activity is comprised of a number of interrelated aspects and that any claims of direct causal relationships require designs where other factors can be controlled for or ruled out. One issue with the study however is that due to a lack of any standardised definition of ritual, the types of 'ritual activities' examined include cross country running, social poker, playing drums and training in capoeira (a Brazilian martial art). Many of these choices are difficult to justify as ritual activities since they have clear instrumental goals and would be more accurately represented as games or athletic training.

Figure 1-19: Conceptual framework of extreme rituals as a social technology



Source: (Fischer & Xygalatas, 2014: 353)

Returning to dysphoric rituals role in group bonding, Fischer & Xygalatas (2014: 352) provide a speculative account that links feelings of moral transgressions with motivation to perform dysphoric rituals and posit that “this experienced pain, driven by perceptions of moral failure is associated with a motivation to donate more to the in-group”. They further suggest that “the apparent supernatural powers of the high-ordeal participants may create a sense of awe, leading to greater supernatural beliefs” which when coupled with “empathetic reactions” (Konvalinka et al., 2011; Xygalatas et al., 2011) leads to “greater identification with the group of believers and social bonding” (Fischer & Xygalatas, 2014: 352). Fischer and Xygalatas (2014: 353) conclude that these individual level psychological effects collectively motivate the performance and observation of the ritual events, which “in turn reinforces social bonds in the collective and acts as a social reinforcement mechanism”, producing a “psycho-biological” feedback loop that operates at both the individual and group level (Fig. 1.20).

Although this model is plausible there is currently very little empirical evidence from actual ritual events for the specific motivations identified or the connections between individual and collective level processes. Furthermore, the previously discussed findings of Van Bunderen & Bastian (2014) suggest that pain has a moral cleansing effect and can reduce generosity towards partners which contradicts the prediction that enduring pain will serve to motivate increased donations to the in-group through “perceptions of moral failure” (Fischer & Xygalatas, 2014: 352).

However, a recent finding that does offer support for Fischer & Xygalatas' (2014) interpretation concerns the so-called *martyrdom effect* detailed in Olivola & Shafir (2013). This paper reports the results of five experiments that demonstrated that, contrary to most theories of motivation and behavior, “the prospect of enduring pain and exerting effort for a prosocial cause can promote contributions to the cause... willingness to contribute to a charitable or collective cause increases when the contribution process is expected to be painful and effortful” (Olivola & Shafir, 2013: 91). Crucially, the martyrdom effect was found to apply across different contexts with people self-reporting a willingness to donate more in a hypothetical condition when they anticipated having to suffer to raise money for a charity (Experiments 1a and 1b) and also demonstrating the same actual behaviour when playing a public goods game that involved real money and required actual pain to increase donations (immersing their hand in ice water) (Experiment 2). The study also demonstrated that the martyrdom effect was “not the result of an attribute substitution strategy (whereby people use the amount of pain and effort involved in fundraising to determine donation worthiness)” (Experiment 3) and that the effect was mediated by perceptions of

meaningfulness (Experiment 4) and moderated by the nature of the prosocial cause with the effect being strongest when directed towards causes associated with human suffering (Experiment 5). Olivola & Shafir (2013) suggest that the prospect of enduring pain generates stronger meaning to their contributions and to the experience of contributing, which in turn motivates higher prosocial contributions. The consistency of the results observed (Table 1.4) across the different experimental conditions increases confidence in the validity of the findings. A crucial point that distinguishes the martyrdom effect from existing dissonance accounts, is that in each case participants reported the amount they wanted to contribute “*before* they experienced any pain-effort associated with the contribution process” and the effect did not vary according to the amount of pain-effort involved or anticipated (Olivola & Shafir, 2013: 102).

Though the findings of Olivola and Shafir (2013) appear to align with the predictions of Fischer and Xygalatas (2014) regarding increased prosocial behaviour with the performance of painful-effortful activities, the mediating and moderating pathways they identify differ. This means further data is required to make any firm judgment on whether the models are compatible. Olivola and Shafir’s emphasis on the important mediating role of a sense of meaningfulness in converting the pain-effort exerted to prosocial behaviour echoes a connection emphasised in Modes theory in relation to imagistic rituals ability to induce self-reflection on their meaning and generate self-defining memories (Whitehouse & Lanman, 2014; Whitehouse, 2004).

Table 1-4: Difference in contributions between effortful and easy tasks

	Relative increase in % contributions	Effect size <i>d</i>	Sample size <i>N</i>	Charity/Cause
Exp. 1a (endurance run vs. picnic)	72	0.70	96	Asian tsunami
Exp. 1b (endurance run vs. picnic)	59	0.51	96	Hurricane Katrina
Exp. 2 (cold pressor vs. control)	31	0.70	36	Public pool (Collective)
Exp. 4 (endurance run vs. picnic)	213	0.61	357	Victims of war and genocide
Exp. 5 (30h fast vs. picnic)	93	0.58	76	Starving children in poor countries
Average [total]	94	0.62	661	

Note: In all comparisons the mean amount contributed was greater in the painful-effortful condition than in the easy-enjoyable one.

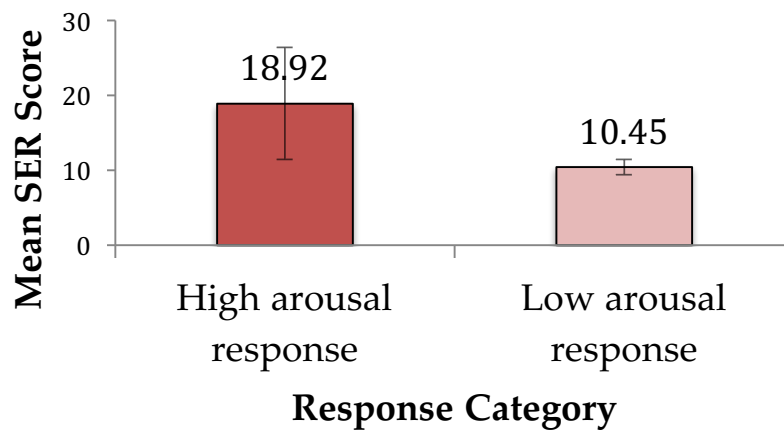
Source: Table adapted from (Olivola & Shafir, 2013: 101)

This aspect of dysphoric rituals has been examined experimentally through the use of a novel artificial ritual methodology (Richert et al., 2005). This study offers a pioneering example of the second category of experimental studies of ritual- those conducted under controlled conditions. Artificial rituals have been presented verbally to examine intuitions about ritual features and causal efficacy (Legare & Souza, 2012) but they have also, in a limited number of studies, involved the actual performance of artificial rituals by groups of participants. In such cases, the studies have then sought to measure the effects that different variations in the ritual setting or structure produce. Richert et al.'s (2005) study employed this design and created two artificial rituals: the first was conducted outdoor in a wood to simulate an ecologically valid ritual setting and the second employed a sound stage, where various lighting and audio features be controlled more precisely. The purpose of the study was to examine whether rituals with dysphoric elements that increased arousal would produce a greater volume and depth of reflection as measured during post-event interviews.

In the first experiment, participants were led by an experimenter in large groups of 14 and 15 people in a performance of a series of ritual actions which included walking across a field while “the leader shook a rattle”, ‘washing’ their hands in rotten leaves on the ground, breaking a part of a stick and throwing it away, drawing a circle around their feet, and reaching down to touch some fur located in a small hole. During these actions certain features of the ritual were varied in order to create a low and high arousal versions. Specifically, the high arousal version took place at dusk, involved participants being blindfolded for segments and required them to throw their sticks “as hard as possible” into the trees (Richert et al., 2005: 135). Self-reported emotional response to the ritual was collected immediately but in-depth follow up interviews were not scheduled until two and a half months later, in order to give the participants time to reflect on their experience. For analysis, participants were categorised post-hoc according to the strength of their affective response rather than condition due to significant individual variations in response to the conditions. Richert et al. (2005: 136) predicted that “individual emotional reactions [would] predict volume and depth of reflection on the experience” and their verbal responses to questions.

This prediction was confirmed through a coding of the content of the follow up interviews, which revealed that the high arousal group displayed a significantly higher amount of attributions of meanings to various actions and a greater volume and depth of analogical comparisons. These results were then combined into a spontaneous exegetical reflection score and the differences observed are shown below (Fig. 1.20).

Figure 1-20: Depth of reflection to a high-arousal artificial ritual

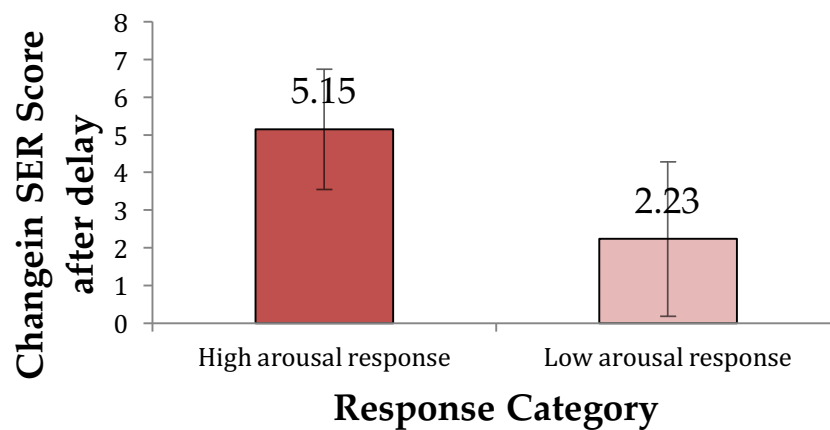


Note: Error bars represent standard deviations.

Source: Adapted from data presented in (Richert et al., 2005: 137)

The second experiment employed a different artificial ritual in order to try and provide greater control over the arousal manipulations. Specifically, the ritual sequence was now performed alone according to instructions that were projected on a screen and was conducted indoors on a specially designed sound stage. The ritual sequence involved was also less complex than the previous outdoor experiment and involved manipulating a variety of items arranged on an altar. Participants were asked to draw a triangle around a necklace and recite a chant, dip their finger in oil and rub it on their forehead and throat, and place their hand inside a box to retrieve a stone. The variance in arousal this time was generated by the volume and intensity of ambient music and drumming, the physical proximity and duration of time that “a man covered in cloth resembling fur” was present, the intensity of lighting conditions (including periods of darkness and red lighting in the high arousal condition) and whether the box that the participant needed to place their hand in was opaque.

Figure 1-21: Change in depth of reflection in artificial ritual



Note: Error bars represent standard deviations.

Source: Adapted from data presented in (Richert et al., 2005: 142)

Questionnaires were again used to measure levels of arousal but an additional physiological measure was collected via galvanic skin response. In-depth interviews coded to quantify level of reflection were conducted both immediately after the performance and following a two-month delay. Following the same post-hoc categorising procedures participants were classified into low and high arousal responses groups and their change in reflective scores were calculated, with the hypothesis being that those “who had a stronger emotional reaction to the ritual” would also demonstrate “a greater increase in their volume and depth of reflection on the ritual between their initial interview and the follow-up interview” (Richert et al., 2005: 143). This prediction was supported as participants who experienced a high arousal response displayed a significantly greater volume of reflection, attribution of meanings and volume and depth of analogic comparisons (combined into an SER scale).

In addition to providing some preliminary evidence for one of the key predictions of the modes theory this study provided an important methodological contribution as it

established that participation in one-off artificial rituals, even under a minimal group paradigm, could produce significant and detectable physiological and psychological effects on participants. The potential salience of unknown rituals was offered additional support by the results from a pilot experiment in which one hundred students observed a single ritual performance and listened to an oral presentation of doctrinal details (2004: 84). The students were tested to examine if there was enough variation to run a full study but it was found that recall of the ritual structure seven weeks later was “more or less perfect” in all students, while the transmitted doctrinal details by comparison had decayed rapidly.

Although these studies provide a good illustration of the future potential for artificial ritual methodologies, the experiments themselves suffer from a number of methodological and sampling issues. For instance, it is puzzling why there are no significant differences in outcomes between the conditions, given that manipulation checks indicated that there was a significant difference in the intended direction in the affective responses they generated. The issue might be related to insufficient sample sizes to detect smaller effect sizes, as the first experiment had only 15 participants per condition and ran only a single group for each of the experimental procedures. Methodologically, this was unlikely to provide experimenters with adequate time to standardise procedures and also meant that there was no recourse for dealing with unanticipated contamination in the sample (e.g. participants knowing each other previously) or the gender disparity and unequal samples in the post-hoc conditions. Furthermore, unequal samples can make the results of t-tests unreliable (de Winter, 2013) and as the significant differences reported are all at the edge of significance ($p=.05$) this decreases confidence in the validity of the findings. The second ‘altar’

experiment addresses some of these concerns but removes the collective element of the dysphoric experience. And again post-hoc affect response groups are used for the analysis, despite manipulation checks indicating significant differences in affective response between the experimental conditions. The differences between the conditions are not reported but it seems likely that they failed to reach significance, which seems to run counter to initial predictions.

Following Richert et al.'s (2005) pioneering experiments, there have been some further attempts to experimentally explore reactions to rituals, however almost unanimously they have employed abstract representations of rituals. For instance, Y. I. Russell et al. (2014) used "the five-disk Tower of Hanoi game" and an invented "Bear God" game that employed the same rules but a distinctive visual and verbal framing that explicitly references rituals, gods, and shamans. The study examined the differential impact of euphoric and dysphoric mood on analogic reasoning and presented the Bear God game "as a proxy for religious ritual" (Y. I. Russell et al., 2014: 1). Euphoric and dysphoric moods were generated through participants watching 10-minute long videos of either a comedy excerpt (euphoric condition) or a realistic depiction of a nuclear war (dysphoric condition). The results indicated that "dysphoria lead to better performance in ... analogical reasoning", specifically for individuals in 'expert' conditions (those who played the Tower of Hanoi task before being introduced to the Bear God game) those who had watched the dysphoric video outperformed those who had seen the euphoric video. This result adds support to the findings reported by Richert et al. (2005) and also lends credence to the case for "differential cognitive roles in religion" for euphoric and dysphoric arousal (Y. I. Russell

et al. 2011) . Yet, the authors also recognised that there are problems in construing a game as a stand in for ritual, especially since it lacked any real sense of “sacral symbolism” (Y. I. Russell et al., 2014: 3). This does not mean that the results are irrelevant to the role that dysphoria can play in ritual settings but it would be premature to assume that the findings are directly applicable to ritual performances.

Another recent experimental study that adopts a similar broad labelling of an artificial ritual is Legare & Wen's (2014) study addressing “the effects of ritual on the development of social cognition”. In this study, children in a day-care setting were assigned to a novel social group then each group participated in necklace making task that was either “scripted, synchronous” and “demonstrated by a group leader” (ritual condition) or “non-scripted” and “supervised [but not demonstrated] by a group leader” (control condition). The results showed that on multiple converging measures participation in the ritual condition increased children’s feelings of in-group affiliation. Specifically, they reported that children in the ritual had higher in-group identity fusion scores, made more affiliative choices and had greater expectations for being included by their in-group. These findings deserve further exploration but they seem to be highly consistent with the predictions of Modes theory (Whitehouse & Lanman, 2014).

However, there is a potential issue with overstating the parallels between the necklace making task and *ritual* as in both the non-ritual and ritual conditions in the study there is a potential instrumental purpose to the activity, no real sense of uncertain authorship, and crucially little symbolic content. This makes the task difficult to justify as an illustration of a “socially shared, conventional and synchronous ritual” (Legare & Wen, 2014), unless all

that is required for ritual is rules, synchrony, and repetitive behaviours. If we adopt such a broad definition then almost everything can be conceived of as a ritual including football matches, chess games, cooking, listening to lectures, and so on. Thus, it seems more productive to regard this study as addressing the effects of certain aspects of ritualised behaviours rather than speaking to ritual more broadly. Furthermore, an alternative mundane explanation for the effects observed that does not seem to have been controlled for is that they derive from the additional attention provided by an authority figure in the ritual condition. Further work is clearly needed to disentangle precisely the mechanisms driving the cohesion observed.

Yet even with these caveats the findings are still impressive and are complemented by research, conducted by the same research team, on the potential role of ritual cues in encouraging over-imitation in children (Legare & Herrmann, 2013; Whitehouse & Legare, 2012). For instance, another recent experimental study conducted with young children manipulated whether a 'ritual like task' had an instrumental or causally opaque outcome, as determined by whether there was a change brought about at the end of the sequence. Since the sequence was performed with young children a simple activity that involved moving a variety of objects in a prescribed sequence was used as 'the ritual like task', a more appropriate label than that used in the previous study. The results obtained indicated that in the causally opaque condition (i.e. the scenario were nothing changed at the end of the sequence) children tended to imitate the series more rigidly than those in the instrumental condition (Legare & Herrmann, 2013; Whitehouse & Legare, 2012). Whitehouse & Legare propose that these findings offer preliminary evidence for the

existence of an early developing *ritual stance* that promotes over imitation of an action sequence when an instrumental purpose is lacking. This in turn might help account for the striking tendency for young children (as well as adults) to imitate causally irrelevant actions when they appear to have been intentionally performed (Keupp, Behne, & Rakoczy, 2013; McGuigan, Makinson, & Whiten, 2011; Nielsen, Mushin, Tomaselli, & Whiten, 2014; Nielsen & Tomaselli, 2010; Whiten, Horner, & Marshall-Pescini, 2005; Whiten, McGuigan, Marshall-Pescini, & Hopper, 2009).

Although these developmental studies, for obvious ethical reasons, do not address dysphoric rituals they highlight the diverse potential applications of employing an artificial ritual, or at least a 'ritual like task', methodology. The main advantage of this kind of approach is that it enables researchers to select specific features of ritual to explore and as such provides a complementary and more ecologically valid means to test the various connections proposed from the theoretical (Bulbulia & Sosis, 2011; Whitehouse & Lanman, 2014; Whitehouse, 2004) and field experiment data (Fischer & Xygalatas, 2014; Xygalatas et al., 2013b). There are a number of research groups, which are now working with artificial ritual paradigms and as a result it is likely that this method of research will continue to develop in the near future.

Yet it is also evident that there are significant limitations with the methodologies employed in existing studies and dysphoric rituals in particular have not been explored directly in a laboratory context since Richert et al. (2005). Consequently, there seems to be significant scope to build on the foundations established by Richert et al. (2005) and further develop the existing artificial ritual methodologies that have been employed in the field

(Legare & Souza, 2012; Legare & Wen, 2014; Y. I. Russell et al., 2014; Whitehouse & Legare, 2012). To this end one of the methodological goals of this thesis is to design and implement an improved artificial ritual experiment that directly manipulates the level of dysphoric arousal. Dysphoric arousal in accordance with the previous discussions in this literature review is defined as referring to the broad category of negative affect as presented in the left hand side of the circumplex model of affect (Fig 1.10). However, it is important to emphasise that what distinguishes dysphoria from negative affect more generally is the accompaniment of higher arousal/activation and as such it does not incorporate negative affective states with low activation such as sadness, depression and boredom but rather stress, nervousness, fear, tension, and discomfort. The newly developed methodology and results from a pilot study in the UK and a full-scale experimental study conducted in Japan are reported in-depth in Chapter 3 of this thesis.

1.11 Summary of the Current Thesis

The preceding literature review has surveyed research from a diverse range of fields including studies of group dynamics, memory formation, cross-cultural psychology, cognitive anthropology, and a variety of ritual theories from CSR. This final section will now briefly outline the current thesis and how it seeks to explore dysphoric ritual practices and the connections between key theories identified in the preceding sections, particularly Modes theory (Atkinson & Whitehouse, 2011; Whitehouse & Lanman, 2014; Whitehouse, 2004), costly signalling (Bulbulia & Sosis, 2011; Henrich, 2009), and identity fusion (Gómez, Brooks, et al., 2011; Swann et al., 2012).

There is a growing experimental literature utilising both field studies and controlled laboratory experiments to explore the motivations and consequences of high arousal rituals. To date, the majority of this work has focused on dysphoric ritual events, but there has still only been a small number of studies and no replications of findings by independent research teams. Currently, there is preliminary evidence, primarily from field work studies, linking participation in collective rituals to enhanced prosociality (Fischer et al., 2013), and these effects appear to be amplified by participating or observing dysphoric rituals (Fischer et al., 2014; Xygalatas et al., 2013b). Such results align with recent experimental work in psychology demonstrating the affiliative effects of shared pain (Bastian, Jetten, & Ferris, 2014) and parallel research in CSR on religion and prosociality from priming and correlational studies (Henrich, 2009; Henrich et al., 2006; Norenzayan et al., 2013; Norenzayan & Shariff, 2008; Shariff & Norenzayan, 2007; Shariff et al., 2015; Weeden & Kurzban, 2013; Xygalatas, 2013a).

This thesis will contribute to this literature by exploring the effects of dysphoric ritual events in the novel context of martial arts groups (Chapter 2), conducting new artificial ritual experiments that manipulate affect and arousal during the performance of collective ritual sequences (Chapter 3), and provide an independent conceptual replication and expansion of recently reported findings linking dysphoric ritual performances at religious festivals with elevated levels of prosociality and inclusive self-identification (Chapter 4). Additionally the role of euphoric rituals (Haidt, Patrick Seder, & Kesebir, 2008) and positive subjective interpretations of dysphoric rituals will be considered in accordance with recent

studies that have sought to address this neglected area of research (Fischer et al., 2014; Y. I. Russell et al., 2011, 2014; Xygalatas, 2014).

The precise nature of the affiliative bonds hypothesised to be generated in collective dysphoric rituals has also become a central theoretical issue for the influential Modes theory (Whitehouse & Lanman, 2014), which proposes that imagistic high arousal rituals generate shared self-defining experiences and, in turn, intense localised ties, characteristic of the newly identified mechanism of identity fusion. Fused individuals, rather than displaying any loss of personal identity or perception of their in-group as homogenous and interchangeable, are motivated by a strong personal identity and a sense of personal agency and regard members of their group “as unique and... irreplaceable members of a larger family” (Gómez et al., 2011: 919). Identity fusion has been operationalized into validated measurement tools (Gómez, Brooks, et al., 2011; Jimenez et al., 2015; Swann et al., 2009) and so the posited connection with high arousal rituals is an empirically tractable issue. Examining this hypothesis could also help to clarify issues associated with parochial altruism, as fusion has been repeatedly shown to be predictive of pro-group behaviour, including willingness to fight and die for the group (Swann, Gómez, Huici, et al., 2010; Swann et al., 2014, 2009). Hence, there is also a need to establish whether the theoretical claim is supported (Swann et al., 2012) that fusion can also be channelled into non-hostile, extreme helping behaviors if group norms determine.

The extent to which the affiliative and prosocial reactions to dysphoric rituals are moderated by pre-existing relationships between performers and observers and mediated by the significance and self-defining qualities attributed to the event, a sense of familial

connection, or perceptions of the pain-effort required, also remain open questions. There is some initial evidence that physiological empathetic responses (Xygalatas et al., 2011) and levels of fatigue (Fischer et al., 2014) experienced by observers are related to their closeness with performers but the boundaries of these effects remain poorly defined. Additionally, it is unclear whether observers who have participated in previous collective dysphoric rituals in the past can come to develop a sense of shared identity or psychological kinship through observation, or whether the visceral connection of simultaneous participation is required. These issues will primarily be addressed by the field studies at firewalking events reported in Chapter 4.

Cognitive dissonance is known to be capable of motivating prosocial and affiliative behaviours following group-relevant dysphoric experiences but most of the research on this has focused on dysphoric initiation events (Aronson & Mills, 1959; Gerard & Mathewson, 1966; Harmon-Jones & Harmon-Jones, 2007). Whether dissonance reduction is a crucial motivation for performers of collective dysphoric ritual events more generally deserves further attention. Another recently developed theory that addresses dysphoric initiation events is *Automatic accrual theory* (Cimino, 2011, 2013). This theory predicts that a group's automatically accrued benefits, such as its status and public facilities, will be positively correlated with the severity of its initiation procedures. At present only survey data from North America (Cimino, 2011, 2013) and selected ethnographic examples have been provided to support this connection. There is thus a need for further cross-cultural tests of this association, as well as the various paradigmatic features that are proposed to distinguish dysphoric initiation events.

The above summary identifies a diverse range of issues that remain unresolved about collective dysphoric rituals and their consequences. However, the underlying thread that unites them is that they are all related to proposed affiliative processes and their downstream effects on prosociality. Consequently, this thesis aims to investigate these processes across a range of contexts, testing the core research questions laid out in the introduction and identifying core mediators and moderators for consistently observed effects. To achieve this, the following chapters report the results from a combination of online surveys (Chapter 2), two controlled laboratory experiments employing a novel artificial ritual methodology (Chapter 3), and field studies of actual dysphoric ritual events performed in Japan (Chapter 4). Ritual, even under restrictive definitions, is a vast topic and entails such a broad and diverse range of phenomena that no single study, regardless of ambition, could adequately address all of the aspects associated with the phenomena. The strategy employed for this thesis therefore takes the opposite approach; it is concerned specifically with exploring highly arousing dysphoric rituals, and their associated psychological impact on group bonding and prosociality. Although other categories of ritual, particularly those that are highly arousing but invoke euphoria, will be addressed overall throughout the thesis the emphasis remains primarily on providing evidence relevant to the wide range of open questions surrounding collective dysphoric rituals and their effects on observers and performers. The approach of this thesis aligns with the methodological trend in CSR to try and fractionate explanations of religious phenomena into manageable chunks and follows recent calls to attempt to combine field and laboratory based investigations (Cohen, 2007; Fischer & Xygalatas, 2014; Xygalatas, 2012).

CHAPTER 2: BRAZILIAN JIU JITSU SURVEY

This chapter reports the results of a large online cross-cultural survey of Brazilian Jiu Jitsu (BJJ) martial art practitioners that was designed to explore whether the level of dysphoric arousal involved in their belt promotion events impacted individuals' feelings of affiliation with their particular school and their willingness to endorse pro-group behaviour. Validated measures of identity fusion and categorical group identification were compared to assess which construct(s) offered more predictive power in relation to the costly pro-group behaviour recorded. The chapter begins with a short summary to provide relevant contextual information about Brazilian Jiu Jitsu, which is a martial art practiced across the world which originated in Brazil, and then outlines the specific rationale for the online questionnaire study and the specific hypotheses being addressed. Following this the actual study is presented; first, details of recruitment and the methodology employed in the online questionnaire construction are provided; second, the results from the survey are reported and analysed; third, the results are discussed in relation to the existing theoretical literature and previous studies.

2.1 Background of Brazilian Jiu Jitsu

Brazilian Jiu Jitsu (BJJ) is a grappling based martial arts system that developed in Brazil as an offshoot from Judo (Peligro, 2003), however it is now practiced across the world. During the 1990s, the early successes of BJJ fighters in the mixed martial arts (MMA) tournament known as the *Ultimate Fighting Championship* (UFC), the largest and most successful MMA association in the world (Watanabe, 2015), dramatically increased the

fame and popularity of the style. BJJ is a popular component of modern MMA training and is well known because of this. Only a minority of BJJ practitioners compete in MMA tournaments however the majority instead compete in grappling tournaments that do not permit striking or do not actively compete outside of training. Due to the emphasis placed on competition and non-compliant sparring in training, BJJ is typically classified as both a martial art and a combat sport.

BJJ schools vary in size, ranging from a few individuals training together in a basement to large academies with multiple affiliates and hundreds of students. Training is conducted in halls and gyms with padded matting on the floor and a typical training session consists of a warm-up, guided practice of specific techniques, and a concluding sparring session. The atmosphere in BJJ schools tends to be significantly less formal than that found in other traditional martial arts; with sparring sessions typically initiated by performing a fist bump and training sessions concluding with all group members individually hugging each other as a sign of camaraderie. Progression through the art is marked via a belt based grading system, similar to that found in Judo or Karate systems (Shun, 1998), however progression is much slower than is typical in similar arts. Judo practitioners, for instance, can usually expect to receive their first belt within six months of beginning to train (Official Judo Rulebook 2008), whereas for BJJ practitioners the first earned belt is typically received after three years of training (Kavanagh, 2013).

Figure 2-1: A belt whipping gauntlet during a BJJ promotion event



Source: *Used with permission of Alessandro Vini (photographed)*

An instructor decides upon the timing of a promotion and belts can be awarded either during formal grading events, or less commonly, as part of normal training sessions. As belt progression is slow, grading events are often significant occasions within a BJJ community and can be highly emotional for those being awarded a higher belt. The precise practices associated with promotions vary between schools but within the community there is a common practice referred to popularly as a belt whipping gauntlet. This practice involves members lining up to form a tunnel along which the student being promoted is required to walk (or run) while the other members whip them repeatedly with their untied belts (Fig. 2.1). Generally, during such events the torso and legs are targeted rather than the face to avoid causing serious injury. The severity of the whipping, whether it is performed with or without a training jacket, and the duration, varies between schools but the practice often results in significant bruising and welts, which recipients now occasionally share on social media (Fig. 2.2).

Figure 2-2: Bruises & welts received during belt whipping gauntlets



Source: *Used with permission of Força Combat Academy and BJJ Eastern Europe*

As BJJ has grown in popularity, particularly in the US, the practice of belt-whipping gauntlets has become increasingly controversial and there are now a number of schools and associations that actively discourage the practice (Huni, 2014). Moreover, online debates about the relevance of the practice in the modern BJJ community often become heated; with those who belong to groups with belt whipping gauntlets defending them as an important bonding experience and those critical deriding them as a form of hazing that harms the image of the sport (Jiu Jitsu Style, 2012). One recent article has also claimed that the practice does not originate from Brazil, as is widely believed, and instead was invented in the US just two decades ago during the 1990s. The account provided in this article describes how a practitioner returning from military training decided to import a hazing ritual (BJJ Heroes, 2015). Although this is unconfirmed, it is certainly plausible that a recently invented tradition could have been imbued with exaggerated historical claims. This is a well-documented practice that has been demonstrated with many iconic cultural practices and rituals across the world (Ranger & Hobsbawm, 1983; Vlastos, 1998). The actual history of

the practice however does not invalidate how the majority of practitioners perceive the history.

Belt whipping gauntlets are not the only painful activity involved with belt promotions, some school's events involve receiving successive throws or slams by the senior members of the club or being required to spar continuously to the point of exhaustion. All of these practices are occasionally discussed with reference to hazing practices, but it is important to note that they are never used as an initial barrier to joining a training group. Indeed, it can be years before members of a particular school become aware of their group's particular promotional practices. Hence, although they may be a sign of deepening commitment they do not represent a prototypical costly initiation.

2.2 Rationale for Study

The current author trained in BJJ for over five years at a university club in London and served as the president for the club²⁶. As a result, I had strong connections across the UK BJJ community and was aware of popular BJJ blogs, podcasts, and forums. I also observed first-hand the significant variation in promotion practices at different BJJ schools and the strong reactions, both positive and negative, that belt-whipping gauntlets invoked. This dramatic variation in promotional practices appeared to offer a unique opportunity to explore in a non-religious context whether high arousal, low frequency dysphoric rituals result in stronger relational group bonds, potentially via identity fusion. BJJ schools may at first seem to involve too little investment to provide a suitable setting to test the predictions of Modes theory, however many BJJ schools promote close, familial connection amongst

²⁶ The present author also served as the president of the University BJJ club for three years.

members and inspire significant levels of commitment amongst members. From my experiences, being a BJJ practitioner is a core part of many practitioners' identity with many making significant personal and financial sacrifices to train regularly. One indication of the importance that can be attached to BJJ by practitioners is the tendency for them to use images with their BJJ training uniforms as their portraits on social media. Although there are likely many factors motivating this decision, including signalling of athletic ability, it does also demonstrate that even when the purpose of training is recreational it can occupy a central part of an individual's projected self-image. Indeed, although there are a small proportion of people for whom BJJ is a professional pursuit but for the majority of practitioners it is a recreational practice that requires significant levels of commitment to make progress.

As mentioned above, belt whipping gauntlets and other dysphoric promotional practices do not serve as barriers to entry, nor are they commonly understood as voluntary displays of commitment. They do function as a rite-of-passage (Van Gennep, 1960; Winslow, 1999), in that they symbolise that an individual has progressed to the next level of ability and seniority. They also contain most of the prototypical features associated with imagistic dysphoric ritual practice (Whitehouse, 2004) in that they are high arousal events, rarely performed and involve repetitive behaviours centred on a demonstration of dysphoric endurance. There is little explicit knowledge about the origins of promotion practices except for vague associations with previous lineages of instructors in Brazil (Humphrey & Laidlaw, 1994), and the reasons for performing certain promotional activities are generally not officially prescribed. Observing discussions at promotion events and in

online discussions also makes it evident that the instrumental purpose or rationale behind practices are often individually generated through personal reflection (Whitehouse & Legare, 2012).

Due to the perceived similarities between BJJ promotional practices and imagistic dysphoric rituals I constructed an online survey to examine whether individuals who reported more dysphoric promotional experiences might, in line with recent predictions of Modes theory (Whitehouse & Lanman, 2014), display higher levels of identity fusion with their BJJ school, more relational bonds with other members, and be willing to endorse and perform more costly pro-group behaviors. A secondary goal of the study was to gather reliable information about the prevalence of belt-whipping gauntlets, and other promotional activities, amongst BJJ groups, as there was no such published information available.

Hypothesis

Drawing on the research literature summarised in Chapter 1, the following specific hypotheses were developed:

H1: Individuals who have participated in belt whipping gauntlets will display a) a greater level of identity fusion with their chosen BJJ school and b) stronger relational bonds with other members of the school.

H2: More dysphoric promotion experiences will be associated with higher levels of identity fusion, when controlling for categorical group identification (Swann et al., 2012; Whitehouse & Lanman, 2014).

H3: The greater the level of dysphoric arousal the less frequent the belt promotion rituals will be (Atkinson & Whitehouse, 2011; Whitehouse, 2004).

H4: Higher levels of identity fusion with the relevant BJJ club will a) be associated with a greater willingness to perform costly sacrifices for the club, when controlling for categorical group identification (Swann, Gómez, Huici, et al., 2010; Swann et al., 2014, 2012) and b) these relationships will be mediated by relational bonds and the significance attached to the experience.

H5: More dysphoric promotion experiences will be associated with a) stronger relational bonds with other group members and b) a greater willingness to engage in costly sacrifice for the group (Swann et al., 2012; Whitehouse & Lanman, 2014). Furthermore, the perception of having endured greater pain and effort in promotion experiences will be associated with greater prosocial contributions towards the ingroup (Bastian, Jetten, & Ferris, 2014; Olivola & Shafir, 2013).

H6: Respondents who endured more dysphoric promotion experiences will regard their groups as having a higher status and greater solidarity (Aronson & Mills, 1959; Cimino, 2013; Kamau, 2013).

H7: More dysphoric experiences will be associated with greater reflection on the experience and attributions of traditional, rather than instrumental purposes (Clegg & Legare, 2015; Richert et al., 2005).

2.3 Methods

Ethics

The data collected from respondents was stored anonymously and all participants were provided with a study information form, required to complete a detailed consent sheet, and provided with a debriefing message at the end of the study²⁷. All procedures for the study were in compliance with the School of Anthropology and Museum Ethnography Research Ethics Committee and received approval (Ref No: SAME/CUREC1A/12-28).

Recruitment Methods & Sample

Respondents were recruited over a six-week period via notices posted on the most popular English speaking online BJJ forums, online message boards, and through word-of-mouth. An interview conducted for a popular BJJ podcast (thefightworkspodcast.com) and the creation of a dedicated website to promote the study (www.bjjsurveys.com) led to a significant amount of interest which was further driven up by members on the BJJ sub-forum of reddit.com. After the exclusion of individuals who had no personal experience with promotion events ($N=165$) and low quality responses ($N=63$), the total sample collected was $N=605$, comprised of 95.1% Males and 4.9% Females²⁸. North Americans accounted for 51.3% of responses; Western Europeans for 14.1% and the remaining 34.6% were widely dispersed across over 30 countries. The age range was from 15 to 64 years ago ($M=31.20$, $SD=7.06$). Time training ranged from less than six months to twenty years ($M=4.57$,

²⁷ A longer debriefing was also made available via email, upon request.

²⁸ A ratio that is typical, in the author's experience, to most BJJ schools and academies.

$SD=3.49$). Due to partial drop out and skipped questions not all respondents answered every question and hence N is reported for each analysis independently.

Questionnaire Content

An online survey was constructed using *Qualtrics* questionnaire software and hosted online. The survey was accessed via a link provided with the online advertisements and took on average of 25 minutes to complete. The full questionnaire is provided in Appendix I but a summary of sections and measures included is provided on the following page (Table 2.1). Overall the questionnaire was designed to collect detailed information about belt promotion practices, measure the individual's subjective response to the experience, and to examine the extent and quality of their group affiliation through measures adopted from Yuki (2003) and validated identity fusion (Gómez, Brooks, et al., 2011) and group identification scales (Mael & Ashforth, 1992).

Three measures were included to examine the respondent's willingness to engage in *costly* sacrifices for their BJJ school. An important point here is that the term *costly* is used in accordance with the terminology found in the identity fusion literature (Swann & Buhrmester, 2015; Swann, Gómez, Huici, et al., 2010) and as such is used to refer to acts that are *personally costly* rather than acts which entail substantial suffering or extreme loss. The term sacrifice was considered to be appropriate since all of the measures used entail the respondent indicating a willingness to give up something (their time, their safety, a potential financial reward) for their group.

Table 2-1: BJJ belt promotions questionnaire content

	SECTION	SAMPLE QUESTIONS & SCALES	MEASURE
	General Training Information	Total Years Training	Years
		Time Training per Week	Hours
		Level of Seniority	Belt
		Group Affiliation	Category
		Competition Experience	Various
		Motivations	Rankings
	Promotion Experience	Time since last attended promotion	Months
		Time since last personal promotion	Months
		Belt whipping experiences	Various
		Throwing experiences	Various
		Sparring experiences	Various
		Positive affective response: how enjoyable?	6-point scale
		Negative affective response: how unpleasant?	6-point scale
		Activation: how intense?	6-point scale
		Importance: how valuable? How meaningful?	6-point scale
		Causal Opacity: how understandable?	6-point scale
		Pain: how painful?	6-point scale
		Reflection on meaning	Open Ended
	Relationship with BJJ School & other members	Pictorial measure of identity fusion (Swann et al., 2009)	6-point scale
		7 item verbal identity fusion scale (Gómez, Brooks, et al., 2011)	6-point scale
		6 item categorical identification scale ²⁹ (Mael & Ashforth, 1992)	6-point scale
		2 item measure of group status (Yuki 2003)	6-point scale
		2 item measure of group entitativity (Yuki 2003)	6-point scale
		DV: 2 item measure of loyalty/willingness to sacrifice (Silver & Brewer, 1997): 1. If my BJJ school really needed me, I would be willing to donate my free time to it. 2. If my BJJ school were threatened, I would be willing to risk my life fighting to defend it.	6-point scale
		DV: 3 item measure of relational bonds – Sociometric bonds (Yuki, 2003) 1. I think all the members of my BJJ school are somehow personally connected to me. 2. I think all the members of my BJJ school are somehow personally connected to each other. 3. I know the personality differences among members of my BJJ school.	6-point scale
	Personality Traits	16 item personal individualistic/collectivist orientation scale (Triandis & Gelfand, 1998)	6-point scale
	Demographics	Sex, Age, Ethnicity, Residence, Native Language, etc.	Various
	Prize Lottery	DV: Voluntary donation opportunity to BJJ school	£0-20

The first two measures were self-reported measures of loyalty (Silver & Brewer, 1997)

that asked the respondents to indicate their agreement with a statement that suggested they

²⁹ This scale was selected as it was previously demonstrated to provide the strongest predictive power for group sacrifice measures, against a selection of other existing group identification scales (Gómez, Brooks, et al., 2011; Swann et al., 2009).

would be willing to 1) donate their time to their BJJ school if it was needed and 2) risk their life if their school was threatened. Both of these were measured using a 6-point scale that ranged from strongly disagree to strongly agree. The third measure was behavioural and asked respondents, who voluntarily chose to participate in a lottery to win £20 (approx. \$30), to designate if they would like to send any of their reward to their school if they were selected as the winner.

Respondents were informed that they could keep all of the money, donate all of it to their school, or split the amount in any ratio they preferred and were also reassured that any amount donated would be sent anonymously to remove potential status-seeking behaviour. Although a relatively low hypothetical payment, participation was entirely voluntary and to opt in to the lottery required extra effort on the part of the respondent and thus it seems likely that any allocations made to the school would be based on genuine motivations to donate rather than demand characteristics of the questionnaire. The introductory text before the donation choice reinforced to respondents that all answers provided would be stored anonymously and that the winners of the lottery would be sent the relevant funds privately via PayPal, using the email address provided.

Potentially confounding demographic and training experience information were also collected, along with a 16 item cross-cultural personality measure that addresses idiocentric and allocentric personality traits (Triandis & Gelfand 1998). The intention here was to help to control for some of the variation inevitable from such a diverse sample. It was also hoped that the variety of measures would help to avoid respondents from guessing the hypothesis through the questions presented.

The sections were presented in the sequence indicated in Appendix I, but within each section, where it did not interfere with branching logic or general survey flow, item presentation was randomised to reduce order effects. Exceptions were made for the identity fusion and group identification scales which recommend against altering the order of presentation³⁰.

2.4 Results

All data analysis reported below was conducted using IBM SPSS Statistical software (Version 22) except for the Confirmatory Factor Analysis which was conducted using the IBM SPSS AMOS (version 19) plugin.

Data Preparation

- Group Status

Two measures of perceived group status from Yuki (2003) were collected using a 6-point scale. The motivation for including this measure was to ensure that predictions deriving from dissonance and signalling theories, that groups which involved more severe initiations or displays of commitment would be considered as higher status. The status measures displayed a weak correlation between items ($r=.216$, $p<.001$) and could not be combined into a viable scale ($\alpha =.501$)³¹. Hence, in the relevant section they are treated independently.

- Group Entitativity

³⁰ Recommended via personal correspondence with Prof. Bill Swann the primary designer of the identity fusion scale.

³¹ This is likely due to the second measure being reverse coded, which can result in different response styles.

Two measures of perceived entitativity, also from Yuki (2003), were collected on a 6-point scale ranging. These items were included to provide an independent measure of group homogeneity, a factor which is central to categorical group identification. The *entitativity* measures displayed a moderate correlation ($r=.600$, $p<.000$) and an acceptable reliability when combined into a scale ($\alpha=.750$).

- Relational Bonds & Sociometric Knowledge

Three measures from Yuki (2003) designed to specifically capture relational bonds and sociometric knowledge about other group members were collected on a 6-point scale. The motivation for including this measure was due to the central role that relational bonds represent in identity fusion theory. The brevity of Yuki's items also eclipsed various alternative relational scales. A reliability analysis of the three measures revealed that one (*I know the personality differences among members of my BJJ school*) had a low corrected item-total correlation ($r=.250$) and severely reduced the reliability of a combined scale. As a result, the three measures were split into a single item measure of *sociometric knowledge* and the remaining 2-items combined into a scale of *relational bonds*. The *relational bonds* items displayed moderate correlation ($r=.76$) and their combined scale had a high reliability ($\alpha=.861$).

- Positive & Negative Promotion Experiences

Seven newly created items were used to measure subjective response to promotion experiences. Specifically, participants rated on a 6-point scale the extent to which they considered their grading experience to be 1) *enjoyable*, 2) *understandable*, 3) *meaningful*, 4) *intense*, 5) *painful*, 6) *unpleasant*, and 7) *valuable*. These items were selected to capture

variables of core theoretical interest to Modes theory predictions and thus included simple positive (1) and negative valenced measures (6), a specific item addressing pain (6), a proxy for activation (4), an indicator of value (7), and items addressing causal opacity (2) and self-reflection (3). It was additionally anticipated that these items would coalesce around positive (1,2,3) and negative (4,5,6) responses and thus serve as broader proxies for euphoric and dysphoric arousal. The presentation of the seven items was randomised for each questionnaire to reduce order effects.

As the items were newly generated a principal components analysis (PCA) was conducted to assess whether the measures were really addressing related latent constructs. The suitability of a PCA was assessed prior to analysis and inspection of the correlation matrix showed that all variables had at least one correlation coefficient greater than .3. As anticipated, the *understandable* item had a large number of similar sized correlations and no individual correlation greater than .5 and was thus excluded from further analysis. The overall Kaiser-Meyer-Olkin (KMO) measure of the remaining 6 items was .73, with individual KMO measures all greater than .6, classifications of 'middling' to 'meritorious' according to Kaiser & Rice (1974). Furthermore, Bartlett's test of sphericity was statistically significant ($p < .001$), indicating that the data was likely factorizable.

The resulting PCA revealed two components that had eigenvalues greater than one, which explained 39.65% and 29.64% of the total variance, respectively. Visual inspection of the scree plot indicated that two components should be retained (Cattell, 1966). In addition, a two-component solution met the interpretability criterion and as such, two components were retained for the final analysis. The two-component solution explained 69.28% of the

total variance with a Varimax orthogonal rotation being employed to aid interpretability. The interpretation of the data was consistent with the responses the items were intended to measure with strong loadings of positive items on Component 1 and negative items on Component 2³². Component loadings and communalities of the rotated solution are presented in Table 2.2. The proposed components also displayed moderate to high reliabilities when combined into scales, see summary lines in Table 2.2.

Table 2-2: Rotated factor loadings for response to belt promotion events

Item	Rotated Factor Loadings	
	Positive (Euphoric)	Negative (Dysphoric)
Valuable	.89	-
Meaningful	.88	-
Enjoyable	.83	-
Painful	-	.88
Unpleasant	-	.80
Intense	-	.78
Eigenvalues	2.42	2.04
% of variance	40.28	33.96
Cronbach's α	.85	.75

Component loadings over 0.6 are highlighted. Component loadings <0.4 not shown.

Four additional measures intended to provide independent measures of the severity of belt whipping experiences were collected through a series of follow up questions to respondents who indicated that they had experienced belt whipping gauntlets ($N=320$). These independent measures addressed the *duration of belt whipping events* (measured on a 4-point ordinal scale, ranging from less than a minute to over 10 minutes), the *number of participants involved* (measured on an 8-point ordinal scale, ranging from a single person to

³² It may be noted here that the label euphoric is overstating the implication of the three positive measures, while conversely negative may be less accurate than dysphoric since it generates the unwarranted assumption that the experience was considered to be negative.

more than 30 people), and whether the whipping took place with or without the protection of a training jacket (coded as a binary category).

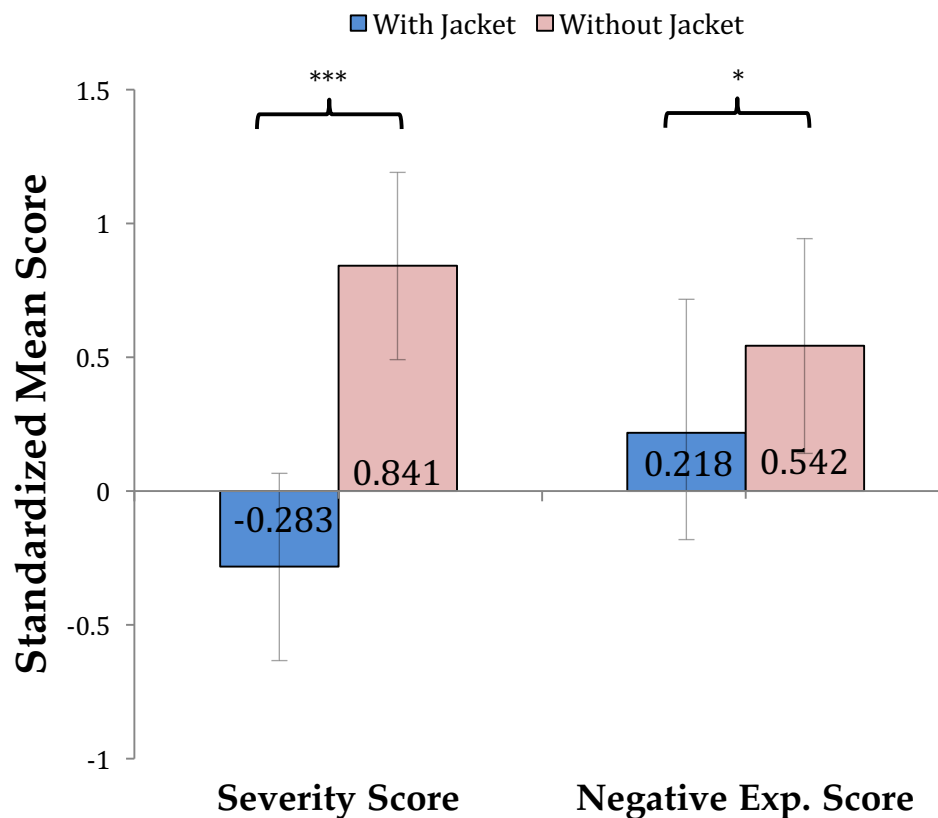
Table 2-3: Spearman's correlations for objective measures of severity

Variable	Duration	Participants	Negative Exp.	Positive Exp.
Duration of Whipping	-	.211***	.307***	.013
Participants in Whipping	.211***	-	.184***	.019
Combined Severity	.643***	.593***	.356***	-.003

*** = $p < .001$, shading added to emphasise significant combined 'severity' measure correlations

The two ordinal measures of duration and number of participants displayed a weak positive correlation ($r_s = .211$, $p < .0001$). Furthermore, as anticipated both measures correlated with the negative experience measure and displayed no statistical relationship with the positive experience measure (Table 2.3). The two measures were consequently standardised and summed to provide a combined *severity* score, which was found to correlate slightly better ($r_s = .356$, $p < .0001$) with the negative experience measure than either of the individual measures. The measure taken in regards to whether or not a jacket was worn during the whipping was employed as a validity check for both the negative experience measure and the combined *severity* score. Both measures demonstrated the predicted pattern with events without training jackets being rated as significantly more negative: $t(279) = 2.383$, $p = .018$, $d = .285$ and more severe: $t(311) = 10.951$, $p < .001$, $d = 1.241$. There was a stronger discrepancy observed with the severity score however this was not surprising given that the measures used to construct the severity item were specifically related to belt whipping, whereas the negative experience measure used items with a broader conceptual frame.

Figure 2-3: Belt-whippings performed without jacket are rated more severe



* $p < .05$, *** $p < .001$, Note: Error bars represent standard deviations.

There are a number of methodological objections to combining ordinal scales (Jamieson, 2004) and the difference conceptually between *duration* and *amount of participants*, together with the weak correlations observed between the two measures, raise concerns of the validity of the *severity* construct. Moreover, the combined measure did possess conceptual validity in that belt whipping events which are both longer and involve more participants are likely to be more *severe*.

As the measure only related to belt whipping experiences it was unsuitable to apply in any analyses that focused on comparative samples across categories. The primary use of this measure in the following analyses was thus to serve as an additional, independent validity check in subgroup analyses specifically examining the belt whipping subgroup. It

was not intended to be used as a core measure for analysis due to methodological issues with combining distinctive ordinal scales.

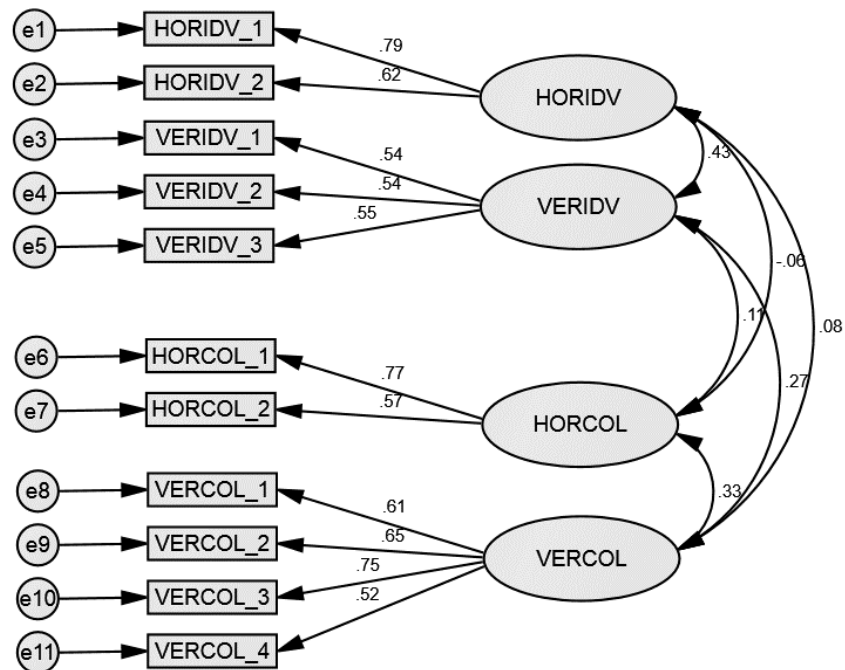
- Idiocentric/Allocentric Personality Traits

Triandis' 16 item individualism/collectivism (IDV/COL) scale is a measure intended to capture individual personality traits of idiocentrism and allocentrism and was included to control for pre-existing personality traits that could influence tendencies to form bonds. The sixteen items included in the scale can be referenced in Appendix I, Section 4. The scale is based on a four factor structure that distinguishes between traits for: i) vertical individualism (VERIDV), ii) horizontal individualism (HORIDV), iii) vertical collectivism (VERCOL), and iv) horizontal collectivism (HORCOL). The measure is widely used in cross cultural psychology and a number of previous studies have provided support for the configural equivalence (J.-S. Chiou, 2001) and metric equivalence of the scale in cross cultural (Gouveia, Clemente, & Espinosa, 2003; Li & Aksoy, 2007; Soh & Leong, 2002) and multi-generational (Guo, Schwartz, & McCabe, 2008) samples.

Yet, critical reviews of the scale also note that certain items display low or unexpected loadings across samples (Cozma, 2011), and factor correlations vary (Oyserman, Coon, & Kemmelmeier, 2002). Therefore, a confirmatory factor analysis (CFA) was performed to assess the suitability of analysing responses in the current sample based on the standard four factor model. From the final sample of 605 respondents, 560 completed the IDV/COL scale and within this data each of the 16 items had less than 1% missing data. A Little's 'Missing Completely at Random' (MCAR) test also confirmed that the missing values appeared to be randomly distributed; $\chi^2(76) = 89.378, p = .140$. As a CFA requires there to be

no missing values in the dataset and there was a limited amount of missing values, substitute scores were calculated using the 'expectation maximization' method. Specifically, values were calculated based on the scores provided for the other three items on each of the relevant subscales.

Figure 2-4: Confirmatory factor analysis for Triandis IDV/COL scale



Note: Standardised values shown.

Following this, a CFA using maximum likelihood estimation (MLE) revealed that the initial factor loading for the expected four factor model had unacceptably low loadings for 4 items: HORIDV3 ($r=.41$), HORIDV4 ($r=.28$), VERIDV4 ($r=.40$) and HORCOL3 ($r=.36$) measures. These items were then excluded and the analysis re-run. In this reduced model only HORCOL4 displayed a loading lower than $<.5$ ($r=.47$) and so it too was excluded and the analysis rerun. The final model retained 11 of the 16 original items, all of which loaded onto their respective factors with a correlation greater than $>.5$ (Fig.2.4). Modification indices identified no items that could be co-varied in accordance with the latent factor

structure and so no further changes were made. Using this model, goodness of fit indices indicated an acceptable model fit with $GFI=.975$, $AGFI=.956$, $CFI=.954$, $RMSEA=.046$, $SRMR=.058$. In contrast, the chi-square value for the overall model indicated a poor fit, $\chi^2(38)=82.359$, $p<.001$, but this measure is known to be unreliable for large sample sizes (Kline, 1998).

Table 2-4: Rotated factor loadings for Triandis IDV/COL scale

Item	Rotated Factor Loadings			
	Horizontal Individualism	Vertical Individualism	Horizontal Collectivism	Vertical Collectivism
HORIDV_3	.732	-	-	-
HORIDV_2	.688	-	-	-
HORIDV_1	.678	-	-	-
HORIDV_4	.478	-	-	-
VERIDV_2	-	.818	-	-
VERIDV_3	-	.682	-	-
VERIDV_4	-	.531	-	-
VERIDV_1	-	.491	-	-
HORCOL_2	-	-	.689	-
HORCOL_1	-	-	.689	-
HORCOL_4	-	-	.651	-
HORCOL_3	-	-	.513	-
VERCOL_2	-	-	-	.778
VERCOL_1	-	-	-	.770
VERCOL_4	-	-	-	.692
VERCOL_3	-	-	-	.666
Eigenvalues	2.303	1.380	1.504	2.838
% of variance	14.395	8.625	9.401	17.735
Cronbach's α	.603	.582	.603	.728

Component loadings over 0.7 and items included in final sub-scales are highlighted. Component loadings <0.4 not shown.

Due to the relatively low loadings, particularly for the vertical individualism factor, an additional exploratory PCA was conducted. The suitability of a PCA was assessed prior to analysis and inspection of the correlation matrix showed that all variables had at least one

correlation coefficient greater than .3. The overall Kaiser-Meyer-Olkin (KMO) measure was .712, with individual KMO measures all greater than .7, classifications of 'middling' to 'meritorious' according to Kaiser & Rice (1974). Furthermore, Bartlett's Test of Sphericity was statistically significant ($p < .001$), indicating that the data was likely factorizable.

The PCA was constrained to a four factor solution in accordance with the theoretical model and this was supported by visual inspection of the scree plot which indicated a four or five component model should be retained (Cattell, 1966). The resultant four-component solution explained 50.16% of the total variance with an Oblimin oblique rotation with Kaiser normalization being employed. This rotation was chosen as not all of the latent factors are orthogonal in the original theory. Resulting component loadings matched theoretical expectations (Table 2.4) but there were four items displaying low loadings of $< .6$.

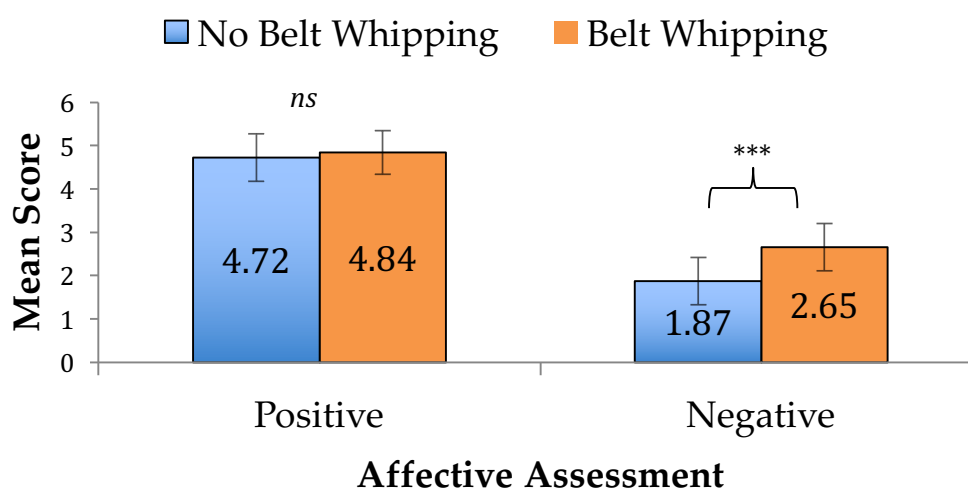
Following both analyses, three items were identified in both as being problematic (specifically, HORIDV4, VERIDV4 and HORCOL3) and as a result these were excluded from analysis. The remaining 13 items were combined to create 4 sub-scales: Horizontal Individualism (HORIDV, 3-items, $\alpha = .613$), Vertical Individualism (VERIDV, 2-items, $\alpha = .560$), Horizontal Collectivism (HORCOL, 3-items, $\alpha = .608$), Vertical Collectivism (VERCOL, 4-items, $\alpha = .728$). The reliability of these sub-scales was lower than anticipated but the structure was retained due to the converging analyses reported above.

Hypothesis 1 (Belt Whipping = More Fusion, More Relational Bonds)

Approximately half of the sample (52.9%) attended BJJ schools which practice belt whipping gauntlets and 96.6% of this group had an experience of the gauntlet at their current school. Using the positive and negative measures of grading experiences as

outcomes, individuals who had experienced belt whippings were compared with those who had not using between groups t-test comparisons.

Figure 2-5: Subjective experience of belt-promotion events



Note: Error bars represent standard deviations.

There was no significant difference found between the two groups in terms of their positive assessment of belt promotion events; $t(561) = 1.243$, $p = .214$, but, as expected, the negative measure was significantly higher for the individuals who had experienced belt whippings; $t(556) = 8.185$, $p < .001$, $d = 0.694$. The severity measure was not used as an alternative comparison as it was only relevant to the belt-whipping group. Further comparisons between these groups for potential confounding factors, including: age, sex-ratio, average training hours per week, years training and idiocentric/allocentric personality traits, revealed no significant differences in the two groups (all $p > .35$).

Table 2-5: Descriptive statistics for H1 variables

Variable	Skew (SE)	Kurt. (SE)	α	M (SD)
Pictorial Identity Fusion	-.271 (.102)	-.959 (.204)	-	3.630 (1.110)
Verbal Identity Fusion	-.144 (.103)	-.341 (.206)	.889	3.834 (1.066)
Group Identification	-.592 (.102)	.420 (.204)	.790	4.28 (.927)
Positive Experience	-.985 (.103)	.255 (.206)	.849	4.790 (.049)
Negative Experience	.866 (.103)	.128 (.206)	.750	2.278 (.050)

Consequently, the groups were compared using bonferroni corrected comparisons on the two outcomes of interest, specifically a) identity fusion, and b) relational bonds with other members. For identity fusion, both the pictorial measure and the 7-item verbal scale were examined, along with the 6-item group identification measure to provide a categorical bonds comparison (descriptive statistics for all variables provided in Table 2.5).

Visual inspection of histograms for the three variables revealed moderate non-parametric distributions but due to the large sample size t-test comparisons were employed³³ (Lumley, Diehr, Emerson, & Chen, 2002). Counter to predictions, no significant difference was found between the groups for a) the fusion or identification measures (all $p>0.35$) or for b) the *relational bonds* measure; $t(563)=.930$, $p=.353$.

In summary, those who endured belt whipping gauntlets as part of their promotions reported the events to be significantly more negative than those who performed other activities. But counter to expectations this difference was not correlated with any significant difference between the groups in terms of levels of identity fusion, group identification, or reported *relational bonds* with other group members. Potential interactive effects between belt whipping experience and the positive/negative measures were further investigated in H2.

Hypothesis 2 (More Dysphoric = Higher Identity Fusion)

A paired t-test on the full sample found that promotion events were reported as being significantly more positive; $M=4.799$, $SD=1.142$, than negative; $M=2.282$, $SD=1.183$,

³³ To provide converging measures, comparisons were additionally replicated using non-parametric Mann-Whitney U tests and all of the patterns reported remained consistent.

experiences: $t(541)=36.956$, $p<.001$, $d=1.587^{34}$. Moreover, this relationship was replicated in both the belt whipping and non-belt whipping samples independently, although in the latter the size of the difference was slightly reduced: Positive- $M=4.875$, $SD=1.103$ vs. Negative- $M=2.649$, $SD=1.160$, $t(260)=23.037$, $p<.001$, $d=1.351$. In the full sample, positive ratings were also not found to be rated orthogonally to negative ratings; $N=563$, $r_s=.075$, $p=.081$. Examining those with belt whipping experience independently, there was no correlation found between the positive and negative measures amongst those who had some experience of gauntlets; $N=281$, $r_s=-.025$, $p=.676$, and this result was also replicated with the severity measure; $N=288$, $r_s=-.006$, $p=.921$. And for those with no belt whipping experience only a weak *positive* correlation was found between the negative and positive experience measures; $N=260$, $r_s=.137$, $p=.027$.

Table 2-6: Spearman's correlation matrix for dys./eup. experiences and group affiliation

Variable	Positive Experience	Negative Experience	Visual Fusion	Verbal Fusion	Group Identification
Positive Experience	-	.075	.345***	.357***	.252***
Negative Experience	.075	-	.024	.078	.078
Visual Fusion	.345***	.024	-	.605***	.299**
Verbal Fusion	.357***	.078	.605***	-	.568***
Group Identification	.252***	.078	.299***	.568***	-

*= $p<.05$, **= $p<.01$, *** = $p<.001$.

Next, to examine whether more dysphoric promotion experiences were associated with higher levels of identity fusion, the negative experience measure was tested for correlations with the single item pictorial measure of identity fusion and mean scores on the verbal identity fusion scale. Using spearman's rho to correct for moderate skew and kurtosis

³⁴ All d 's reported from within subjects comparisons of means apply corrections for dependence between means, using equation 8 from Morris & DeShon (2002).

(Table 2.6), no significant correlation was found between negative experiences and scores on the pictorial fusion measure; $N=524$, $r_s=.024$, $p=.587$, but for the verbal scale of identity fusion, there was a weak correlation that was just outside of statistical significance; $N=518$, $r_s=.078$, $p=.077$. However, a similar result was observed between negative experiences and the group identification measure; $N=541$, $r_s=.078$, $p=.081$, which implied that any affiliative effect from the negative experience was not specifically related to the identity fusion bonds. To help clarify if either connection was more substantial a partial correlation between the verbal fusion and negative experience measures was run controlling for identification, with bootstrapping (1,000 samples) applied to help adjust for non-parametric distributions and no significant correlation was detected; $r=.045$, $p=.306$, BC_a CI [-.040, .136].

Correlations for the same fusion measures were also examined with the positive experience measure and, counter to predictions, stronger statistically significant correlations were discovered for all three outcomes. Additionally, the strongest correlation ($r_s=.357$) was observed with the verbal identity fusion measure (Table 2.6). Again, a partial correlation was conducted to examine if the connection between verbal fusion and positive experience remained when controlling for group identification, bootstrapping (1,000 samples) was again applied. The results revealed that a significant positive correlation was retained; $r=.259$, $p<.001$, BC_a CI [.179, .355], with only a minor reduction in magnitude. Examining the alternative visual fusion in the same manner a similar correlation was found; $r=.297$, $p<.001$, BC_a CI [.222, .362], increasing confidence in the validity of the finding. In contrast, when controlling for verbal fusion, the correlation with group identification remained significant but decreased from $r_s=.252$, $p<.001$ to less than half of its original

magnitude; $N=520$, $r=.088$, $p<.044$, BC_a CI [.005, .168]. This indicated that the relationship observed between identity fusion and positive experience was more robust than that observed with categorical group identification.

A sub-group analysis examining the belt whipping and non-belt whipping responses independently also provided further support for this interpretation as the correlation between positive experience and group identification, controlling for verbal fusion, was only found in the non-belt whipping sample; $N=251$, $r=.140$, $p=.026$, BC_a CI [.009, .287]. A non-significant relationship was observed in the non-belt whipping sample; $N=266$, $r=.035$, $p=.573$, BC_a CI [-.086, .154]. Conversely, the correlation between verbal fusion and positive experience was replicated in both of the samples independently; non-belt whipping- $N=251$, $r=.194$, $p=.002$, BC_a CI [.005, .168] and also appeared to represent a stronger relationship in the belt whipping subgroup; $N=266$, $r=.327$, $p<.001$, BC_a CI [.005, .168].

To investigate these relationships in greater depth a hierarchical linear regression involving three stages was conducted with mean verbal identity fusion as the outcome measure. In the first step, a binary measure to indicate the presence/absence of belt whipping was entered along with demographic variables (age and sex), and other potential confounding variables: idiocentric/allocentric personality scores; group identification³⁵; average hours training per week; total years training; and months since last promotion experience³⁶. In the second step, subjective positive and negative assessment scores of

³⁵ As entitativity and group status are components of group identification they were not included as separate factors in models including the identification measure.

³⁶ Relational bonds were not included in this model as they were anticipated to be a key part of the mechanism connecting fusion and affective responses to promotion experiences and thus could potentially reduce the magnitude of any connection.

promotion experiences were added to generate the 'full' model, and, finally, in the third step, an additional model with the interactions between the binary belt whipping experience measure and the positive/negative assessment scores were added. An initial correlation analysis found no significant relationship between fusion and sex ($r=-.034$, $p=.239$) or total years training ($r=.001$, $p=.498$) and so these were excluded from the model.

Testing the remaining variables and the model generated for assumptions: first, examination of standardised residuals identified 3 cases as being extreme outliers with values greater than 3 SDs. Further examination of leverage and influence, using Cook's and Mahalanobis distance values, indicated no other problematic cases. Consequently, the three outliers were removed and analyses were rerun. From this model, there was independence of residuals, as assessed by a Durbin-Watson statistic of 2.063. Partial regression plots also indicated suitably linear relationships for all non-categorical variables. There was no indication of multicollinearity as assessed through examination of correlations and VIF and tolerance statistics, except for the positive ($VIF=.099$, $TIF=10.131$) and negative ($VIF=.087$, $TIF=11.524$) measures and their respective interaction terms with belt whipping experience. To aid interpretation the correlation coefficients of the positive and negative experience measures were centred prior to analysis. Finally, examination of the residuals in a histogram, P-P plot, and Normal Q-Q plot indicated an appropriately normal distribution.

Table 2-7: Regression models predicting identity fusion

Variable	Outcome: Identity Fusion (Verbal)					
	STAGE 1		STAGE 2		STAGE 3	
	B (SE)	β	B (SE)	β	B (SE)	β
Constant	-.787 (.435)	-	-.387 (.432)	-	-.355 (.429)	-
Age	-.009 (.006)	-.060	-.011 (.006)	-.069	-.009 (.006)	-.059
Train Hrs.	.023 (.009)	.098***	.019 (.008)	.082*	.019 (.008)	.083*
Last promotion	-.004 (.002)	.003	-.004 (.002)	-.053	-.004 (.002)	-.055
Identification	.679 (.043)	.593***	.557 (.044)	.486***	.563 (.044)	.491***
Hor-Idv.	.074 (.052)	.052	.101 (.051)	.072*	.103 (.051)	.074*
Ver-Idv.	.074 (.046)	.062	.068 (.045)	.058	.063 (.045)	.053
Hor-Col.	.285 (.057)	.189***	.272 (.055)	.181***	.256 (.055)	.170***
Ver-Col.	.042 (.043)	.038	.017 (.042)	.015	.014 (.042)	.013
Belt-Whip	.066 (.076)	.031	.027 (.078)	.013	.034 (.077)	.016
Positive	-	-	.172 (.034)	.187***	-.073 (.099)	-.079
Negative	-	-	.018 (.033)	.020	.228 (.104)	.253*
Pos*Belt Whip	-	-	-	-	.164 (.064)	.274**
Neg*Belt Whip	-	-	-	-	.134 (.065)	-.236*
R^2	.462		.493		.505	
F	41.549***		38.304***		33.777***	

$N=445$, *= $p<.05$, **= $p<.01$, ***= $p<.001$, shading added to emphasis variables with significant coefficients.

The full model, with all of the variables included in stage 2 (Table 2.8) was able to significantly predict a substantial portion of the variance in the verbal identity fusion measure; $R^2=.493$, $F(11,433)=38.304$, $p<.001$. Although this represented a statistically significant improvement over the initial model (including only demographic and potentially confounding variables), the increase in explanatory power was minor; $\Delta R^2=.031$, $F(2,433)=13.208$, $p<.001$. This limits the significance of the interaction analyses reported from stage 3 but considering the level of heterogeneity inherent in a worldwide sample and the number of alternative factors controlled for a minor increase in predictive power is still a potentially important finding.

Adding the interaction terms to the model also only generated a very minor increase in the variance accounted for; $\Delta R^2=.011$, $F(2,431)=4.992$, $p<.001$. But the new model also found a significant interaction between belt whipping and positive experience scores, which superseded the independent contribution of the positive experience scores. Furthermore, a weaker negative interaction was also found with negative experience scores and belt whipping experience. These results, coupled with the previous analyses reported above, suggest that fusion was higher amongst those with belt whipping experience who also had a more positive experience, with the inverse relationship evident amongst those who rated the event more negatively. The consistent, significant contribution of horizontal-collectivist (allocentric) personality scores also suggests that the construct contains parallels with the identity fusion measure or that this is a potential personality level predictor of fusion tendencies.

Hypothesis 3 (More Dysphoric = Less Frequent)

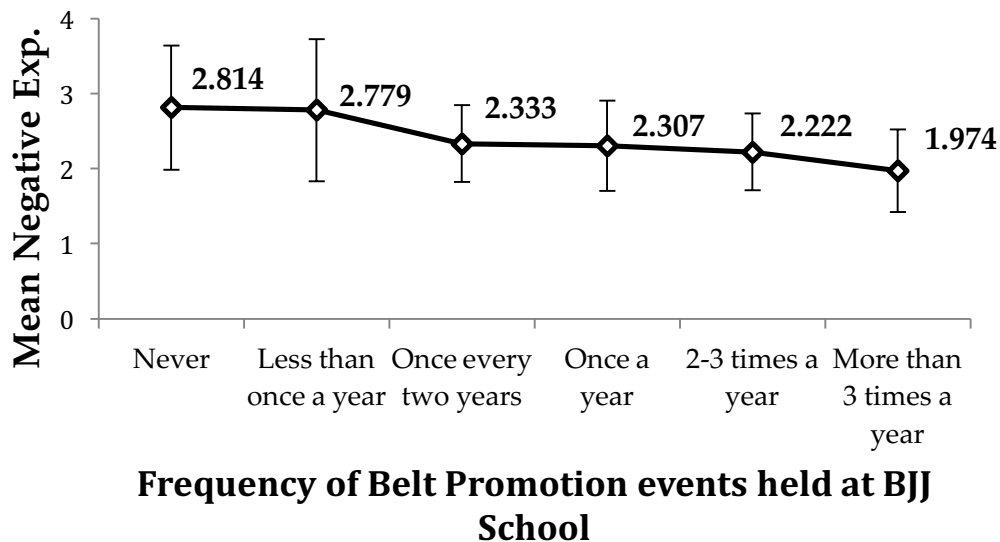
To examine whether more dysphoric experiences were associated with less frequent performance of belt promotion events a simple correlation analysis was conducted between the frequency of promotions³⁷ (measured on a 6 point ranked scale, with higher values representing a higher frequency) and the negative and positive experience measures.

There was no significant relationship found between the level of positive experience and frequency; $N=407$, $r_s=.024$, $p=.627$. However, in accordance with the hypothesis, there was a

³⁷ 200 responses were excluded from analysis due to respondents not answering the question or indicating that the scheduling of promotion events was random.

weak but significant inverse correlation with the negative experience measure; $N=399$, $r_s=-.101$, $p=.044$ (Fig. 2.6).

Figure 2-6: Decrease in dysphoric ratings with increased frequency of promotions



Note: Error bars represent standard deviations.

Hypothesis 4a (Higher Fusion = More Costly Sacrifice³⁸)

Kolmogorov-Smirnov tests for normality indicated non-parametric distributions for the three personally costly sacrifice measures and this was confirmed by visual inspection of histograms and q-q plots (Table 2.8). The donate time measure was severely negatively skewed, due to 74.2% of respondents selecting higher responses (either agree or strongly agree). Conversely, the willingness to risk life measure was positively skewed, with 58% selecting lower responses (either disagree or strongly disagree). The group donation measure had a bimodal distribution with 60.7% of respondents choosing to donate £0 and 27.6% choosing to donate the full £20. Due to these non-parametric distributions, initial correlations were explored using spearman's rho.

³⁸ See the data preparation section for a clarification on the intended definition of costly sacrifices.

Table 2-8: Descriptive statistics for self-reported sacrifice measures & donations

Variable	N	Skew (SE)	Kurt. (SE)	M (SD)
Donate Time	570	-1.213 (.102)	-1.955 (.204)	5.050 (1.008)
Risk Life to Defend	568	-.710 (.103)	-.677 (.205)	2.570 (1.594)
Donate £ to Group	384	-.592 (.102)	-1.328 (.248)	6.63 (8.812)

Initial correlation analyses using the verbal fusion measure, which has previously proven the strongest measure for predicting the endorsement and performance of sacrificial behaviour, displayed moderate correlations with both the time donation; $N=561$, $r_s=.515$, $p<.001$, and the risking life for the group measure; $N=559$, $r_s=.546$, $p<.001$. A weaker but still highly significant correlation was found with the group donation behavioural measure; $N=377$, $r_s=.250$, $p<.001$. These findings were in accordance with predictions however categorical group identification displayed a similar weaker pattern of correlations (Table 2.9).

Table 2-9: Correlations for sacrifice and group affiliation measures

Variable	Donate Time	Risk Life to Defend	Donate to Group
Group Identification	.380***	.336***	.165***
Identity Fusion	.515***	.546***	.250***

*** = $p<.001$; overall group identification – fusion correlation ($N=560$, $r_s=.568$, $p<.001$).

To help distinguish whether the verbal identity fusion measure was able to predict each of the sacrifice outcomes independently and to a greater extent than the group identification measure a series of multiple regression analyses were conducted. There is some debate about the validity of using ordinal data as the dependent variable in regression analysis (Jamieson, 2004), but several studies have demonstrated that ordinal data can return accurate values, including from F tests, if assumptions are met (Carifio & Perla, 2007; Glass, Peckham, & Sanders, 1972). For the following regressions all

assumptions were examined and marginal significance values were treated with additional caution.

H4a Sacrifice Measure 1 (Willingness to Donate Time)

In order to help identify potentially confounding external variables, an initial multiple regression was conducted on willingness to donate time entering the following variables as predictors: demographic factors- age and sex; positive and negative measures; a binary measure to indicate the presence/absence of belt whipping; idiocentric/allocentric personality scores; group identification; average hours training per week; total years training; and months since last promotion experience.

Testing assumptions: there was acceptable independence of residuals, as assessed by a Durbin-Watson statistic of 1.941 and partial regression plots indicated approximate linear relationships with all non-categorical variables. There was also no indication of multicollinearity as assessed through correlation analysis and VIF and tolerance statistics. An examination of standardized residuals from the final model indicated 7 cases as extreme outliers with values over 3 SDs and a further 2 cases demonstrated extreme values of leverage and Mahalanobis distance. These 9 cases were removed and the model was rerun. Finally, a visual examination of the resulting residuals in a histogram, P-P plot, and Normal Q-Q plot indicated a normal distribution.

The model that emerged significantly predicted willingness to donate time; $F(6,530)=7.118$, $p<.001$, $R^2=.119$, but age, sex, belt whipping, total years training, and all idiocentric/allocentric personality scores (except horizontal collectivism) did not contribute significantly to the final model (Table 2.10). Therefore, the only variables retained as

additional predictors for the subsequent regression models were: the positive experience measure, horizontal collectivism (allocentric) scores, average training time per week, and time from last promotion.

Table 2-10: Exploratory regression with willingness to donate time as outcome

Variable	b (SE)	Standardized β	Sig.
Constant	3.582 (.506)	-	.000
Sex	.070 (.197)	.016	.721
Age	-.003 (.006)	-.026	.588
Train Time (Yrs.)	-.000 (.014)	.000	.992
Belt Whipping	-.076 (.084)	-.042	.367
Avg. Train Week (Hrs.)	.039 (.009)	.197***	.000
Last Promotion	-.007 (.003)	-.116*	.013
Hor-Idv.	.040 (.055)	.033	.472
Ver-Idv.	.005 (.048)	.005	.913
Hor-Col.	.236 (.059)	.183***	.000
Ver-Col.	.049 (.045)	.052	.278
Positive	.188 (.036)	.235***	.000
Negative	-.011 (.035)	-.014**	.766

p<.05*, *p*<.001***, shading added to emphasis significant predictors in the model

Based on the exploratory analysis, a two-stage hierarchical regression was conducted on willingness to donate time with average hours training per week, positive experience, horizontal collectivism, time from last promotion, and group identification, entered in the first stage and identity fusion added in the second. All assumptions were re-examined but no violations or additional outliers were detected, so no additional responses were removed.

The model in the first stage significantly predicted willingness to donate time; $F(5,455)=34.843$, $p<.001$, $R^2=.277$, with group identification as the strongest contributor to the model. However, when identity fusion was added in the second stage the model that emerged predicted a greater amount variance, $\Delta R^2=.04$, $F(6,454)=36.190$, $p<.001$, $R^2=.324$, and

identity fusion was the strongest contributing variable. Furthermore, the relative contribution of group identification decreased dramatically, almost halving in magnitude (Table 2.11).

Table 2-11: Hierarchical regression on willingness to donate time

Variable	Outcome: Donate Time			
	STAGE 1		STAGE 2	
	B (SE)	β	B (SE)	β
Constant	2.088 (.295)	-	2.270 (.288)	-
Train Hrs.	.032 (.008)	.158***	.026 (.008)	.129***
Last promotion	-.005 (.002)	-.091*	-.004 (.002)	-.075
Hor-Col.	.205 (.055)	.156***	.128 (.055)	.098*
Positive	.129 (.034)	.159***	.084 (.034)	.103**
Identification	.307 (.042)	.312***	.173 (.047)	.176**
Identity Fusion	-	-	.256 (.046)	.291***
R^2	.277		.324	
F	34.843***		36.190***	

$N=460$, *= $p<.05$, **= $p<.01$, *** = $p<.001$, shading added to emphasis variables with significant coefficients.

H4a- Sacrifice Measure 2 (Willingness to risk life)

As with the previous analysis, an initial multiple regression to identify potentially confounding external variables was conducted on the willingness to risk life measure (Table 2.12). Assumptions of linearity, independence of errors, homoscedasticity and normality of residuals were met and there were no significant outliers or cases displaying undue leverage. The model that emerged significantly predicted willingness to risk life; $F(12,441)=9.579$, $p<.001$, $R^2=.207$. Three of the items that contributed to the model (average hours training per week, positive experience, horizontal collectivism) matched the variables that predicted willingness to donate time but the vertical individualism and negative experience scores also made significant contributions to the model (Table 2.12). Negative

experience scores did not technically fall within significance but given the level of uncertainty and the relevance of the construct it was judged as appropriate to include in the full model.

Table 2-12: Exploratory regression with willingness to risk life as outcome

Variable	b (SE)	Standardized β	Sig.
Constant	-1.721 (.890)	-	.000
Sex	-.096 (.340)	-.012	.778
Age	-.013 (.011)	-.055	.236
Train Time (Yrs.)	-.009 (.023)	-.019	.700
Belt Whipping	.007 (.144)	.002	.963
Avg. Train Week (Hrs.)	.067 (.016)	.190***	.000
Last Promotion	-.001 (.005)	-.015*	.750
Hor-Idv.	-.019 (.094)	-.009	.838
Ver-Idv.	.176 (.082)	.099*	.033
Hor-Col.	.303 (.102)	.135**	.003
Ver-Col.	.109 (.078)	.066	.161
Positive	.345 (.061)	.249***	.000
Negative	-.118 (.061)	.089	.052

p < .05, p < .001***, shading added to emphasis significant predictors in the model*

Based on this exploratory analysis, a two-stage hierarchical regression with willingness to risk life time as the outcome variable was conducted. In the first stage, group identification was entered with the five potential confounding variables and in the second stage, only identity fusion (verbal) was added so that its contribution could be parcelled out. After the model was run all assumptions were re-examined but no violations or additional outliers were detected, so no additional exclusions, or changes to the model were made.

The model in the first stage significantly predicted willingness to donate time; $F(6,482)=22.212$, $p<.001$, $R^2=.217$, with positive experience and group identification as the strongest contributing variables. However, when identity fusion was added in the second

stage; $F(7,481)=36.190$, $p<.001$, $R^2=.353$, the model improved significantly; $\Delta R^2=.136$, $F(1,481)=101.219$, $p<.001$, and group identification no longer contributed significantly to the model. Instead, its contribution was fully mediated by fusion, which demonstrated a substantial correlation coefficient despite the model controlling for the various other potentially confounding factors (Table 2.13).

Table 2-13: Hierarchical multiple regression on willingness to risk life

Variable	Outcome: Donate Time			
	STAGE 1		STAGE 2	
	B (SE)	β	B (SE)	β
Constant	-2.088 (.544)	-	-1.641 (.502)	-
Train Hrs.	.057 (.015)	.162***	.039 (.014)	.110**
Ver-Idv.	.122 (.073)	.070	.043 (.067)	.025
Hor-Col.	.237 (.093)	.108*	.022 (.087)	.010
Positive	.296 (.059)	.212***	.166 (.055)	.119**
Negative	.074 (.055)	.056	.059 (.050)	.044
Identification	.346 (.075)	.203***	-.025 (.077)	-.015
Identity Fusion	-	-	.741 (.074)	.494***
R^2	.217		.353	
F	22.212***		37.457***	

$N=460$, $*=p<.05$, $**=p<.01$, $***=p<.001$, shading added to emphasis variables with significant coefficients.

H4a- Sacrifice Measure 3 (Costly Donation)

Examining responses to the hypothetical donation measure ($N=380$) revealed a clear bimodal distribution caused by 60.8% of respondents donating nothing to the group and 27.4% donating the full £20. Such a distribution is unsuitable for a linear regression model but an exploratory regression was conducted to examine correlations and detect potential outliers. No extreme outliers or cases displaying undue leverage were detected and hence all responses were retained for analysis. However, inspection of standardized residuals confirmed a bimodal distribution was retained and this was unable to be corrected through

transformations. Therefore, a standard hierarchical linear regression was judged as inappropriate to explore the relationship between the potential predictor variables.

Table 2-14: Correlation with donation measure

Variable	r_s	BCa 95%	Sig.
Sex	.131* (.060)	[.004, .250]	.020
Age	.011 (.057)	[-.094, .122]	.839
Train Time (Yrs.)	-.047 (.054)	[-.154, .061]	.404
Belt Whipping	.035 (.055)	[.055, -.079]	.530
Avg. Train Week (Hrs.)	.040 (.055)	[-.063, .144]	.479
Last Promotion	-.040 (.057)	[-.152, .063]	.482
Hor-Idv.	-.086 (.059)	[-.198, .043]	.127
Ver-Idv.	-.015 (.058)	[-.147, .092]	.788
Hor-Col.	.134* (.056)	[.027, .234]	.017
Ver-Col.	.118* (.056)	[.022, .223]	.035
Positive	.146* (.055)	[.030, .245]	.009
Negative	.033 (.059)	[.071, .141]	.470

p < .05*, shading added to emphasis significant correlations

As regression coefficients were likely to be unreliable, simple correlations with donation were examined instead, using spearman's rho with bootstrapping applied (1,000 samples). Four variables displayed a weak correlation with donation; the two allocentric scales, the positive experience measure, and sex (Table 2.14).

In order to maintain a degree of consistency to the analysis with the other sacrifice measures two cumulative odds ordinal logistic regression with proportional odds were conducted after converting the donation variable into an ordinal scale with three levels (£4 or less, £5-£14, £15-20). This three level categorisation was chosen to avoid discarding the 11.2% of responses that fell between the two distribution peaks. The first stage included the potential confounding factors and the second stage was the full model with identity fusion and group identification measures added. The details of the model provided below concern

the full model from stage 2, additional information for the stage 1 model is available in Table 2.15.

In the second stage, in accordance with the preliminary correlation analysis, mean scores for group identification and identity fusion (verbal) were entered along with the two allocentric variables and positive experience scores³⁹. VIF and tolerance statistics indicated no multicollinearity in the data and there were proportional odds, as assessed by a full likelihood ratio test comparing the residual of the fitted location model to a model with varying location parameters, $\chi^2(5)=6.385$, $p=.270$. The deviance, $\chi^2(701)=604.029$, $p=.997$, and pearson, $\chi^2(701)=706.811$, $p=.431$, goodness-of-fit tests indicated that the model was a good fit to the observed data, but the large number of empty cells (66.7%), resulting from the inclusion of four continuous predictors, made these findings unreliable.

Fortunately, confidence in the model was enhanced by the final model, with all 5 variables entered, demonstrating a significantly improved ability to predict donations than the intercept-only model, $\chi^2(5)=35.690$, $p<.001$. However, from the variables entered, in accordance with predictions, only fusion (Table 2.15) demonstrated a statistically significant effect on the model's predictive power; *Wald* $\chi^2(1)=10.951$, $p=.001$. An increase in fusion (measured using a 6-point scale) was associated with an increase in the odds of donating to the group, with an odds ratio of 1.562, 95% CI [1.199, 2.033], *Wald* $\chi^2(1)=10.951$, $p<.001$.

³⁹ Sex was excluded as a predictor due to the small number of relevant female responses in the sample, $N=17$, which would make any findings obtained unreliable.

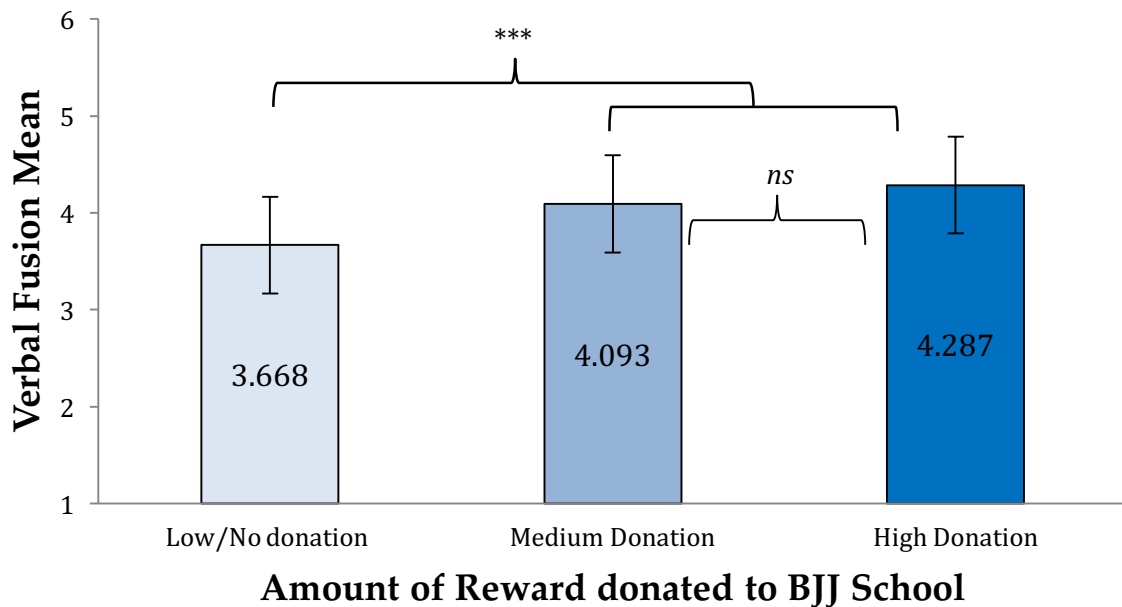
Table 2-15: Hierarchical ordinal regression on willingness to donate

Variable	Outcome: Donate Prize Money			
	STAGE 1		STAGE 2	
	Threshold B (SE)	Wald χ^2	Threshold B (SE)	Wald χ^2
Donations .00	4.328 (.939)	21.235***	4.673 (.990)	22.278***
Donations .01	3.883 (.947)	26.604***	5.275 (.999)	22.914 ***
Ver-Col.	.109 (.121)	.815	.016 (.125)	.017
Hor-Col.	.441 (.163)	7.270**	.296 (.169)	3.07
Positive	.273 (.113)	5.792**	.169 (.115)	2.163
Identification	-	-	.029 (.143)	.042
Identity Fusion	-	-	.446 (.135)	10.951
χ^2	17.453***		35.690***	

*N=359, *= $p<.05$, **= $p<.01$, *** = $p<.001$, shading added to emphasis variables with significant coefficients.*

In summary, from the sample of respondents who chose to participate in the reward lottery, higher levels of identity fusion were associated with greater voluntary donations to their BJJ school. This effect remained significant when controlling for personality factors, positive promotional experience, and categorical group identification.

Figure 2-7: Fusion score for different levels of donations to the group



*$p<.001$ ***, Note: Error bars represent standard deviations*

To examine this relationship further a one-way ANOVA was conducted comparing level of verbal identity fusion between three ordinal donation categories (donated £4 or less, donated £5-£14, donated £15-20). Levene's statistic indicated that there was homogeneity of variance ($p=.92$), and a main effect was found for donation category on level of identity fusion; $F(2,359)=14.382$, $p<.001$, with a significant linear trend; $F(1,376)=28.282$, $p<.001$, indicating that as the amount donated increased the level of fusion (based on the verbal measure) increased proportionately.

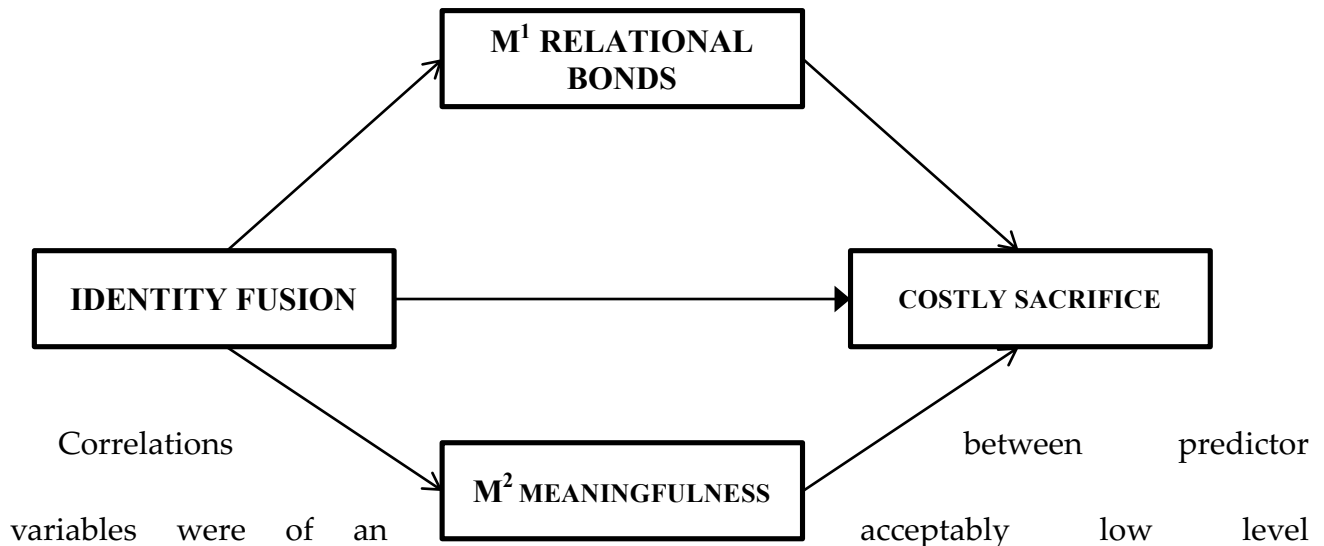
Planned contrasts indicated that those who donated more than £4 had a significantly higher level of fusion than those who donated a lower amount or nothing; $t(376)= 4.583$, $p<.001$, $d=0.473$, but found no significant difference in fusion between those who donated all of their potential winnings with those who donated a medium amount portion; $t(376)=-1.054$, $p=.293$.

Hypothesis 4b (Mediation of effects by Relational bonds & Meaningfulness)

Having confirmed that identity fusion was associated with costly sacrifices on all three measures, the next step was to examine if these connections were mediated by i) relational bonds with other members of the BJJ school, using the combined *relational bonds* measure described in the data preparation section, and ii) how meaningful the promotion experience was judged to be. A series of mediation analysis were conducted using PROCESS (Hayes, 2012) and the modelling recommendations laid out in Hayes (2013). A mediation model with two parallel mediating variables (M_1 = combined *relational bonds* measure and M_2 = single item measure of *meaningfulness*) was constructed to explore these factors as potential

mechanisms underlying the connection between identity fusion and the costly sacrifice outcomes (Hayes, 2013: Model 4).

Figure 2-8: Schematic representation of the two mediator model (Hayes 2013)



addressing concerns that multicollinearity could skew the results (Table 2.16).

Table 2-16: Spearman's correlation matrix for sacrifice measures, fusion and mediators

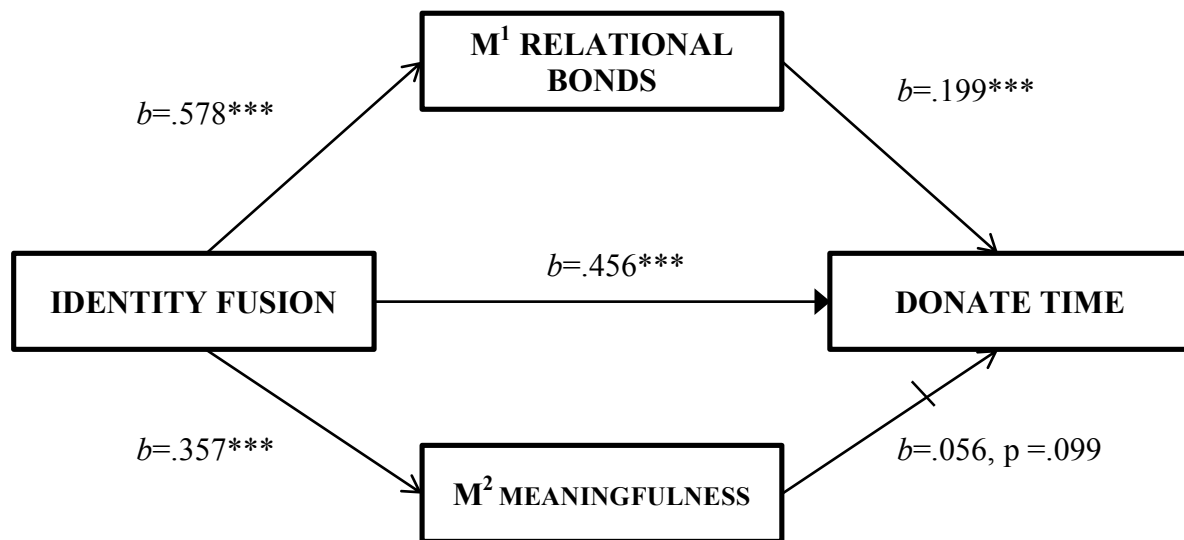
Variable	Identity Fusion	Relational Bonds	Meaningfulness	Donate Time	Risk Life	Donate
Identity Fusion	-	.530***	.316***	.515**	.546**	.250**
Relational Bonds	.530***	-	.206**	.420**	.368**	.163**
Meaningfulness	.316**	.206**	-	.240**	.282**	.118*

p<.01**, = p<.001***

H4b- Sacrifice Measure 1 (Willingness to donate time)

On the first costly sacrifice outcome measure of willingness to donate time there was a significant indirect effect of fusion through *relational bonds*; $b=.115$ ($SE=.028$), BC_a CI [.068, .178], but no significant indirect effect through meaningfulness; $b=.020$ ($SE=.013$), BC_a CI [-.001, .051]. The full model pathways and standardized measures of effect size (Preacher & Kelley, 2011) are reported in Fig 2.9.

Figure 2-9: Relational bonds mediate fusion's effect on donations of time



N=537

Direct effect $b=.320$ (.048), $p < .001$

Indirect effect M¹ $b=.115$ (.028), 95% BCa CI [.068, .176]

Indirect effect M² $b=.020$ (.013), 95% BCa CI [-.001, .051]

Index of Mediation M¹ $\beta=.123$ (.028), 95% BCa CI [.074, .185]

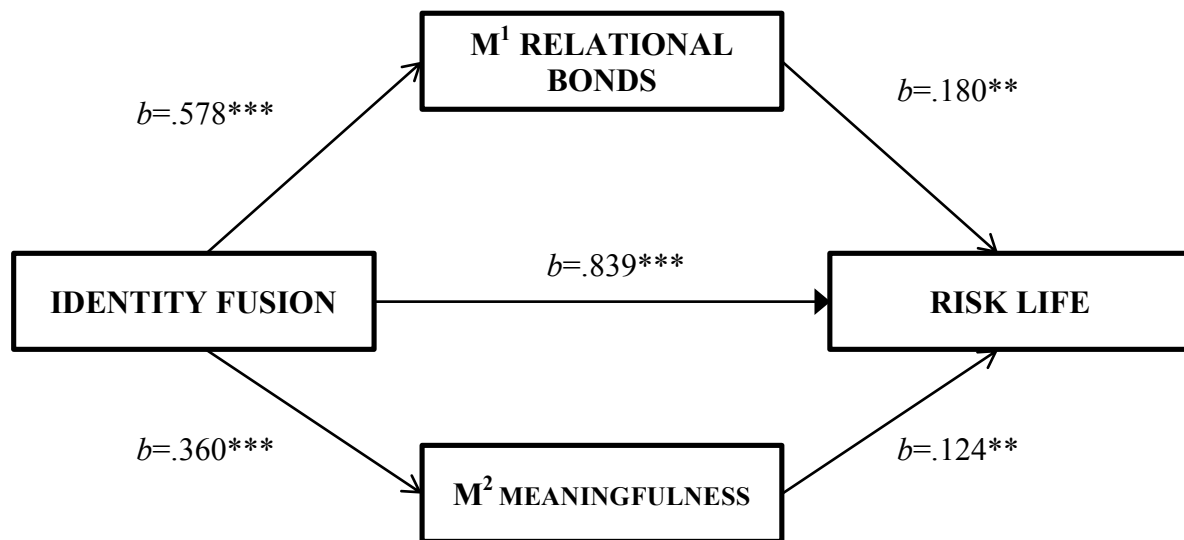
Index of Mediation M² $\beta=.021$ (.014), 95% BCa CI [-.002, .054]

Bootstrapping for Bias Corrected CI conducted with 1,000 Samples

H4b- Sacrifice Measure 2 (Willingness to risk life)

On the second sacrifice outcome measure of willingness to risk life for the BJJ school there was a significant indirect effect of fusion through *relational bonds*; $b=.104$ ($SE=.034$), BCa= CI [.035, .167], and a smaller, but still significant, indirect effect through *meaningfulness*; $b=.045$ ($SE=.018$), BCa CI [.011, .084]. The full model pathways and standardized measures of effect size are reported in Fig 2.10.

Figure 2-10: Relational bonds and meaningfulness mediate fusion's effect on risking life



N=537

Direct effect $b=.690$ (.065), $p < .001$

Indirect effect M^1 $b=.104$ (.035), 95% BCa CI [.035, .173]

Indirect effect M^2 $b=.045$ (.018), 95% BCa CI [.012, .079]

Index of Mediation M^1 $\beta=.069$ (.023), 95% BCa CI [.023, .113]

Index of Mediation M^2 $\beta=.030$ (.012), 95% BCa CI [.008, .053]

Bootstrapping for Bias Corrected CI conducted with 1,000 Samples

H4b- Sacrifice Measure 3 (Costly Donation)

Given the bimodal, non-parametric distribution of the costly donation outcome, it was anticipated that mediation models would be based on inaccurate model parameters. Despite this, an exploratory mediation analysis using the same mediation model was assessed but the analysis was abandoned as no indirect pathways were found to be significant, with all confidence intervals crossing zero.

In summary, mediation analyses of the relationship between the costly sacrifice measures and identity fusion found the predicted mediating relationship with *relational bonds* for the two self-report measures. A weaker and less consistent mediation was detected with ratings of how *meaningful* belt promotion experiences had been. For the

behavioural donation measure, no mediation was detected but the non-parametric distribution of the outcome variable makes this finding unreliable.

Hypothesis 5a (More Dysphoric = Greater Costly Sacrifices)

To test the prediction that more dysphoric experiences would promote a greater willingness to sacrifice for the group, the negative experience measure was examined against the same three costly sacrifice measures reported for H4a.

An initial correlation analysis with the positive and negative measures revealed that the negative measure displayed a weak correlation with only one item, risking life for the group; $N=523$, $r_s=.111$, $p=.011$. Counter to hypotheses, but in line with the findings reported for H4, the positive experience measure displayed significant correlations with all three sacrifice outcomes (Table 2.17).

Table 2-17: Spearman's correlations for sacrifice measures and negative and positive experience

	Donate Time	Risk Life to Defend	Donate £ to Group
Negative Experience	.001	.111*	.043
Positive Experience	.287***	.322***	.176***

$p < .01^{**}$, $= p < .001^{***}$

Building on the analyses conducted in H4, three stage hierarchal multiple regression analyses were conducted to explore the extent to which subjective response to belt promotions was predictive of each of the three sacrifice measures, when controlling for other identified confounding factors (see H4a). In the first stage of the model, previously identified confounding factors were entered; in the second stage, the positive and negative experience measures were added to examine the extent to which they improved the model, and finally, in the third stage, identity fusion, identification, and interaction terms with the

positive/negative experience measures were added to further explore the relationships detected in H4. Group affiliation measures were excluded until the third stage despite being known predictors of the outcomes as these were anticipated to be a part of the pathways through which affective experiences generate affiliation and thus to potentially disguise the relationships between the outcome measure and the positive/negative variables.

H5a- Sacrifice Measure 1 (Willingness to donate time)

Prior to analyses, the 9 extreme outlier/leverage cases identified in H4a were excluded and all assumptions were rechecked. Partial regression plots indicated approximate linear relationships with all new variables and there was no indication of multicollinearity as assessed through correlation analysis and VIF and tolerance statistics, except for the relevant variables and the interaction terms. To aid the interpretation of coefficients in stage 3 all variables were centred prior to analyses. Finally, a visual examination of the residuals in a histogram, P-P plot and Normal Q-Q plot demonstrated a normal distribution for the final model.

The model in stage 2, including known confounds and positive/negative experience measures, significantly predicted willingness to donate time; $F(4,471)=27.828$, $p<.001$, $R^2=.191$, but counter to predictions, it was only the positive experience measure that contributed significantly to the final model. Conversely, no significant relationship was detected with the negative experience measure and the willingness to donate time measure (Table 2.18). Furthermore, the addition of positive and negative experience measures explained only a small additional amount of the variance ($\Delta R^2=.044$) than was already

accounted for in the previous model with average hours training and the horizontal collectivism allocentric personality score. In the third stage, as expected the contribution of group identification and fusion superseded the positive and negative experience measures but there was no significant interaction detected (Table 2.18). There was also a larger increase in predictive power ($\Delta R^2=.111$)

Table 2-18: Hierarchical multiple regression on willingness to donate time

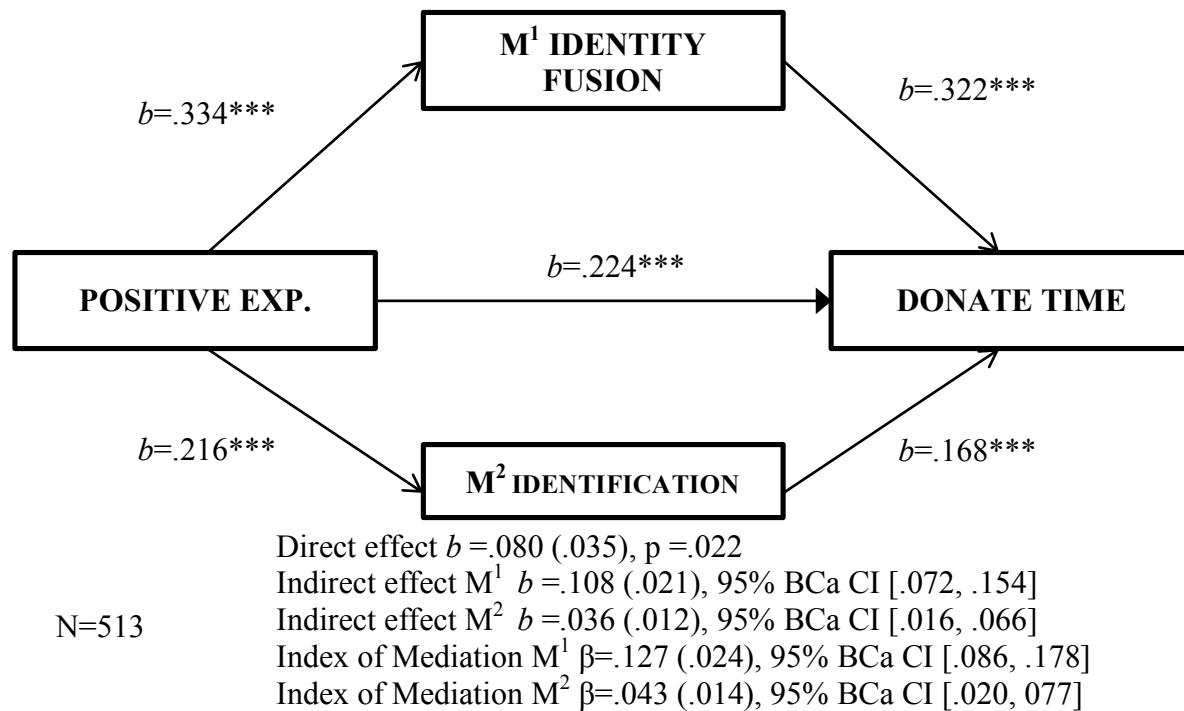
Variable	Outcome: Donate Time					
	STAGE 1		STAGE 2		STAGE 3	
	B (SE)	β	B (SE)	β	B (SE)	β
Constant	3.130 (.250)	-	3.308 (.247)	-	4.056 (.249)	-
Train Hrs.	.051 (.010)	.220***	.048 (.010)	.207**	.032 (.009)	.139***
Hor-Col.	.349 (.053)	.283***	.315 (.052)	.255***	.176 (.052)	.143***
Positive	-	-	.167 (.033)	.211***	.078 (.112)	.099
Negative	-	-	-.034 (.032)	-.046	.051 (.114)	.068
Identification	-	-	-	-	.143 (.046)	.348*
Identity Fusion	-	-	-	-	.296 (.146)	.147**
Fusion*Positive	-	-	-	-	-.002 (.029)	.981
Fusion*Negative	-	-	-	-	-.023 (.028)	.405
R^2	.147		.191		.302	
F	40.892***		27.928***		25.274***	

$N=476$, * = $p < .05$, ** = $p < .01$, *** = $p < .001$, shading added to emphasis variables with significant coefficients.

To explore these results in greater depth, a mediation model was constructed with the verbal identity fusion and group identification measures entered as parallel mediators between positive experience and time donation outcome measure (Fig. 2.11). There was a significant indirect effect of positive experience operating through identity fusion; $b=.108$, ($SE=.021$), and a weaker indirect effect through group identification, $b=.036$ ($SE=.012$). Moreover, the direct effect was more than halved in magnitude when the indirect pathways were included, indicating that these accounted for a substantial portion of the observed

relationship. The full model pathways and standardized measures of effect size are reported in Fig 2.11.

Figure 2-11: Mediation of positive experience effect on donations of time



Bootstrapping for Bias Corrected CI conducted with 1,000 Samples

H5a- Sacrifice Measure 2 (Willingness to risk life)

Next, examining the sacrifice measure of willingness to risk life for BJJ school, another three stage hierarchical linear regression was conducted. In the first stage, the significant confounding variables identified in H4a were entered and then in stage 2, the positive and negative measures were added. Finally, in stage 3 the group affiliation measures of identity fusion and group identification were added along with interaction terms with the positive and negative measures. Assumptions of linearity, independence of errors, homoscedasticity, and normality of residuals were met. There were also no significant outliers or cases displaying undue leverage detected so no additional responses were removed.

Table 2-19: Hierarchical multiple regression on willingness to risk life

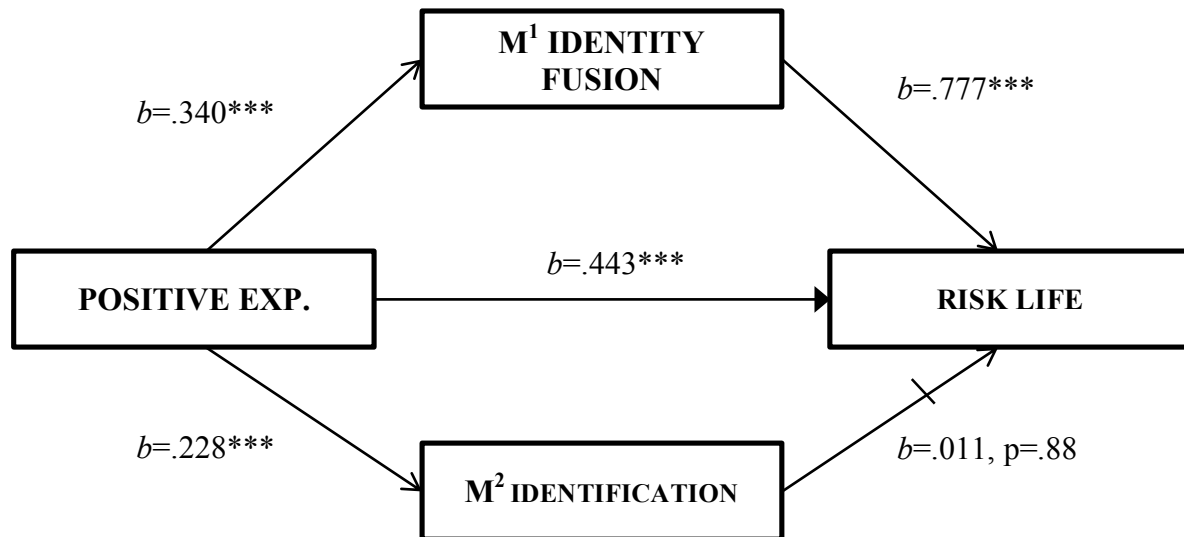
Variable	Outcome: Risk Life					
	STAGE 1		STAGE 2		STAGE 3	
	B (SE)	β	B (SE)	β	B (SE)	β
Constant	.158 (.458)	-	-1.332 (.498)	-	4.422 (1.014)	-
Train Hrs.	.071 (.017)	.180***	.061 (.017)	.153***	.022 (.015)	.056
Hor-Col.	.410 (.097)	.188***	.330 (.094)	.151***	.006 (.087)	.003
Positive	-	-	.360 (.059)	.260***	-.368 (.179)	-.266*
Negative	-	-	.098 (.056)	.074	-.121 (.183)	-.091
Identification	-	-	-	-	.010 (.077)	-.006
Identity Fusion	-	-	-	-	-.054 (.240)	-.036
Fusion*Positive	-	-	-	-	.148 (.046)	.707**
Fusion*Negative	-	-	-	-	.049 (.044)	.177
R^2	.077		.151		.350	
F	20.062***		21.315***		31.955***	

$N=482$, * = $p < .05$, ** = $p < .01$, *** = $p < .001$, shading added to emphasis variables with significant coefficients.

The second stage model, including known confounds and positive/negative measures significantly predicted willingness to risk life; $F(4,478)=21.315$, $p < .001$, $R^2=.151$, representing an increase in explanatory power from the initial confound only model of $\Delta R^2=.074$. Even though this represented only a minor improvement it is worth noting that the only additional contributing variable in stage 2 was the measure of positive experience. Negative experience in contrast made no significant contribution to the model. This was counter to predictions but echoed the findings from the previous analyses (Table 2.19). An intriguing interaction was also observed in the final stage between the verbal identity fusion measure and the positive experience measure. The observed effect superseded the contribution of identity fusion alone which was not a significant predictor in the model and reversed the observed coefficient for the positive experience measure. However, it should be cautioned that despite using centred values substantial levels of collinearity between the interaction

variables and the individual variables persisted and as a result coefficient values may be unreliable.

Figure 2-12: Mediation of positive experience effect on willingness to endorse risking life



$N=513$

Direct effect $b = .176$ (.054), $p = .001$
 Indirect effect M¹ $b = .264$ (.042), 95% BCa CI [.179, .343]
 Indirect effect M² $b = .003$ (.017), 95% BCa CI [-.030, .037]
 Index of Mediation M¹ $\beta = .190$ (.028), 95% BCa CI [.131, .244]
 Index of Mediation M² $\beta = .002$ (.012), 95% BCa CI [-.022, .026]

Bootstrapping for Bias Corrected CI conducted with 1,000 Samples

If the interaction effect observed is valid then it implies that identity fusion and positive experiences were a core factor in determining reported willingness to risk life for the BJJ school as demonstrated by a doubling of the variance ($\Delta R^2 = .201$) accounted for in stage 3 as compared with stage 2.

To further examine these results, a mediation model, following the same structure as with the time donation measure, was constructed with verbal identity fusion and identification measures entered as parallel mediators (Fig. 2.12). On this outcome measure

the significant indirect effect of positive experience through identity fusion was replicated but there was no indirect effect observed through group identification; $b=.003$ ($SE=.017$), BCa CI $[-.030, .037]$. This means that the partial mediation observed ($b=.267$), was entirely accounted for by identity fusion. The full model pathways and standardized measures of effect size are presented in Fig 2.12.

H5a- Sacrifice Measure 3 (Costly Donation)

Due to the bimodal distribution the constructed ordinal scale with three levels (£4 or less, £5-£14, £15-20) was used as the outcome variable in a two stage cumulative odds logistic regression with proportional odds. Following the findings reported in H4, the two allocentric personality measures were entered as potential confounds with the positive and negative experience scores in stage 1. Following this in stage 2 an exploratory model was created that added group identification and verbal identity fusion measures. This stage was included as the distribution of the donation measure meant that a mediation analysis, as used with the other two sacrifice measures, would be unreliable.

For stage 1, VIF and tolerance statistics indicated no multicollinearity in the data and there were proportional odds, as assessed by a full likelihood ratio test comparing the residual of the fitted location model to a model with varying location parameters, $\chi^2(4)=6.510, p=.164$. The deviance, $\chi^2(662)=587.422, p=.983$, and pearson, $\chi^2(662)=668.764, p=.419$, goodness-of-fit tests indicated that the model was a good fit to the observed data. But the large number of empty cells (66%) resulting from the inclusion of four continuous predictors, made these tests unreliable.

However, confidence in the model was improved independently as it was found to predict donations over and above the intercept-only model, $\chi^2(4) = 20.137, p < .001$. An increase in positive experience rating was associated with an increase in the odds of donating to the group, with an odds ratio of 1.284, 95% CI [1.320, 2.187], *Wald* $\chi^2(1) = 4.829, p = .028$. Conversely, there was no significant contribution to the model from the negative experience measure, *Wald* $\chi^2(1) = 2.004, p = .224$. Horizontal collectivist traits were the strongest significant predictor in the model (Table 2.19).

Table 2-20: Hierarchical ordinal regression on willingness to donate

Variable	Outcome: Donate Prize Money			
	STAGE 1		STAGE 2	
	Threshold B (SE)	Wald χ^2	Threshold B (SE)	Wald χ^2
Donations .00	4.443 (.966)	21.142***	4.698 (1.034)	20.637***
Donations .01	5.012 (.974)	26.478***	5.287 (1.043)	25.718***
Ver-Col.	.065 (.121)	.288	-.042 (.128)	.109
Hor-Col.	.464 (.162)	8.162**	.334 (.173)	3.699
Positive	.250 (.114)	4.829*	.377 (.115)	1.578
Negative	.127 (.090)	2.004	.102 (.093)	1.202
Identification	-	-	.011 (.145)	.006
Identity Fusion	-	-	.458 (.135)	11.304**
χ^2			35.690***	

*N=359, * = $p < .05$, ** = $p < .01$, *** = $p < .001$, shading added to emphasis variables with significant coefficients.*

In stage 2, when group affiliation measures were added model assumptions were re-assessed. Again, VIF and tolerance statistics indicated there were no issues of multicollinearity and there were proportional odds, $\chi^2(6) = 6.761, p = .344$. Similarly, the deviance, $\chi^2(674) = 580.801, p = .996$, and pearson, $\chi^2(674) = 683.145, p = .395$, goodness-of-fit tests indicated that the model was a good fit to the observed data. This was also supported by the statistical improvement in predictions over an intercept only model,

$\chi^2(6)=34.670, p<.001$. Unlike the model in stage 1, the only significant predictor of donation was identity fusion; $B=.458$ ($SE=.135$), $Wald \chi^2(1)=11.304, p=.001$.

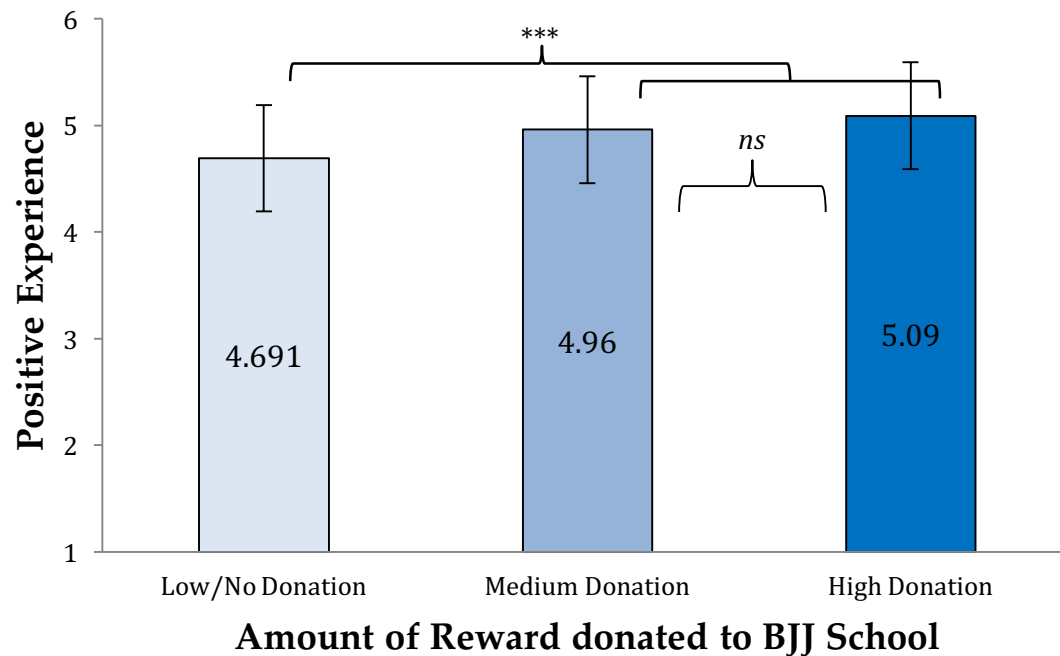
In summary, amongst the respondents who participated in the reward lottery, more positive promotion experiences were associated with higher group donations, whereas negative experiences demonstrated no significant independent effect. However, the observed effect of positive experiences was subsumed when identity fusion was entered into the model. This reaffirms the findings for the previous two sacrifice measures but could not be confirmed directly through mediation analysis due to the bimodal distribution of the outcome.

Interaction terms between fusion and positive experiences were entered into a further exploratory ordinal regression model but severe violations of multicollinearity were encountered that could not be resolved through centering procedures, which meant the model produced was unreliable and thus further analysis was not conducted. Preliminary examination of parameter estimates suggested that the interactions offered no significant independent contribution to the model but did reduce the observed relationship with positive experience to below significance thresholds.

Reversing the model parameters, a one-way ANOVA applying the ordinal donation categories as the grouping variable and positive experience ratings as the outcome was conducted. There was homogeneity of variances, as assessed by Levene's test for equality of variances ($p=.494$) and a significant difference in the mean level of positive experience emerged; $F(2,359)=5.156, p=.006, \eta^2=.028$, with a significant linear trend across categories;

$F(1,359)=11.378$, $p=.002$, indicating that as the amount donated increased, the level of positive experience reported increased proportionately.

Figure 2-13: Positive experience and donations to the group



$p<.001$ ***, Error Bars represent standard deviations

Planned contrasts also revealed that those who donated some amount reported significantly more positive experiences than those who donated nothing; $t(359)= 2.759$, $p=.006$, but there was no statistically significant difference in positive experience scores from those who donated £5-£14 and those who donated £15-£20; $t(359)=.652$, $p=.515$. Post-hoc comparisons using Hochberg's GT, which is robust against differences in group size, also detected a difference between donation groups that was on the margin of significance ($p=.058$).

H5a- More Pain = More Costly Sacrifices

Finally, as there is evidence that pain can serve to enhance in cooperation in social scenarios (Bastian, Jetten, & Ferris, 2014; Olivola & Shafir, 2013), an additional correlation analysis examining the single item pain measure with the three costly sacrifice measures was conducted but no significant correlations were found with any of the measures (all $p > .17$).

Taken collectively the results from all of the analyses conducted for H5a demonstrated that, counter to predictions, the three pro-group sacrifice measures displayed no strong relationship with negative ratings of promotion experiences. Instead, a robust relationship was found between positive ratings of promotion experiences and the sacrifice measures. Furthermore, the connections between positive experience and the costly sacrifice measures were repeatedly found to be partially mediated by identity fusion to a greater extent than with categorical group identification. Additionally, counter to some preliminary findings from recent studies (Bastian, Jetten, & Ferris, 2014; Xygalatas et al., 2013b) pain was not found to be associated with increased prosociality, as operationalised through the three costly sacrifice measures.

Hypothesis 5b (More Dysphoric = Stronger Relational Bonds)

Having addressed the relationship between positive and negative measures and the three costly sacrifice measures, the next hypothesis was that negative experiences would be associated with the formation of stronger relational bonds with other group members who share these experiences. To examine this hypothesis, the connections between positive and negative assessments of promotion experiences and the measures of *relational bonds* and

sociometric knowledge were examined. The prediction was that a more dysphoric promotion experience would promote stronger relational bonds and additionally that there would be a stronger connection with the *relational bonds* measure than the less personal measure of *sociometric knowledge*. To examine this first, a correlation analysis using spearman's rho was conducted due to moderate skew and kurtosis (Table 2.21).

Table 2-21: Descriptive statistics for relational measures

Variable	Skew (SE)	Kurt. (SE)	α	M (SD)
Relational Bonds	-.443 (.103)	-.009 (.205)	.861	3.934 (1.133)
Sociometric Knowledge	-1.015 (.102)	2.189 (.204)	-	4.800 (.887)
Positive Experience	-.985 (.103)	.255 (.205)	.846	4.790 (1.151)
Negative Experience	.866 (.103)	-.128 (.206)	.745	2.228 (1.184)

As predicted, the results demonstrated a positive correlation between negative promotion experiences and *relational bonds* with other group members; $N=524$, $r_s=.097$, $p=.030$. However, the correlation observed was weak and additionally a stronger correlation was found with positive experiences; $r_s=.257$, $p<.001$. Similarly, for the *sociometric knowledge* item, a weak correlation was found with both the negative experience measure; $r_s=.096$, $p=.028$, and the positive experience measure; $r_s=.118$, $p=.006$.

In order to explore these connections further, exploratory multiple regressions were conducted on both the *relational bonds* and *sociometric knowledge* measures to identify potential confounds. Demographic factors- age and sex; the binary measure to indicate the presence/absence of belt whipping; idiocentric/allocentric personality scores; group identification; average hours training per week; total years training, months since last personal promotion, and months since last grading attended were all entered in to exploratory regression models (Table 2.22).

Table 2-22: Exploratory regression of relational bond and sociometric knowledge

Variable	Relational Bonds		Sociometric Knowledge	
	b (SE)	Standardized β	b (SE)	Standardized β
Constant	1.429 (.591)	-	3.728 (.520)	-
Sex	-.016 (.238)	-.003	.068 (.202)	.016
Age	-.013 (.007)	-.082	-.006 (.006)	-.049
Train Time (Yrs.)	-.013 (.017)	-.038	-.001 (.014)	-.004
Belt Whipping	-.012 (.098)	-.005	-.054 (.083)	-.031
Avg. Train Week	.052 (.011)	.215***	.021 (.009)	.108*
Last Promo Attnd.	-.015 (.007)	-.104*	-.011 (.005)	-.119*
Last Pers. Promo	-.000 (.004)	-.116	.002 (.003)	.027
Hor-Idv.	.016 (.067)	.011	.051 (.056)	.044
Ver-Idv.	.109 (.058)	.088	.005 (.049)	.005
Hor-Col.	.305 (.072)	.196***	.160 (.060)	.131**
Ver-Col.	.190 (.055)	.166***	.041 (.046)	.045

$p < .05^*$, $p < .01^{**}$, $p < .001^{***}$, shading added to emphasis significant predictors in the model

The models for both outcomes demonstrated independence of residuals and partial regression plots also revealed approximate linear relationships with all non-categorical variables. There was also no indication of multicollinearity as assessed through correlation analysis and VIF and tolerance statistics. However, although visual examination of residuals through a histogram, P-P plot, and Normal Q-Q plot indicated a normal distribution for the *relational bonds* measure, there was a severely non-normal distribution for the *sociometric knowledge* measure. Attempted transformations were unable to normalise the distribution to an acceptable degree and as a result further analyses of the *sociometric knowledge* measure in a regression model was deemed to be inappropriate.

Focusing then on the *relational bonds* model, an examination of standardized residuals indicated 4 cases as extreme outliers with values over 3 SDs and a further 2 cases demonstrated extreme values of leverage and Mahalanobis distance. These 6 cases were removed and the model was rerun. After exclusions, the model that emerged significantly

predicted *relational bonds*; $F(11,427)=8.599$, $p<.001$, $R^2=.181$, however the only variables that contributed to the model and were retained as additional predictors were the allocentric collectivism scores, average training time per week, and time from last promotion attended.

Based on the exploratory analysis, a three-stage hierarchical regression was constructed with the *relational bonds* measure as the outcome with demographic factors and other significant predictors entered in the first stage, positive and negative experience measures in the second, and identity fusion and group identification added in the third. Fusion and identification were not entered in the initial stage, as a mediating relationship with identity fusion was anticipated based on the results observed in H5a. All assumptions were re-examined prior to the launch of the final model but no violations or additional outliers were detected, so no further changes or exclusions were made.

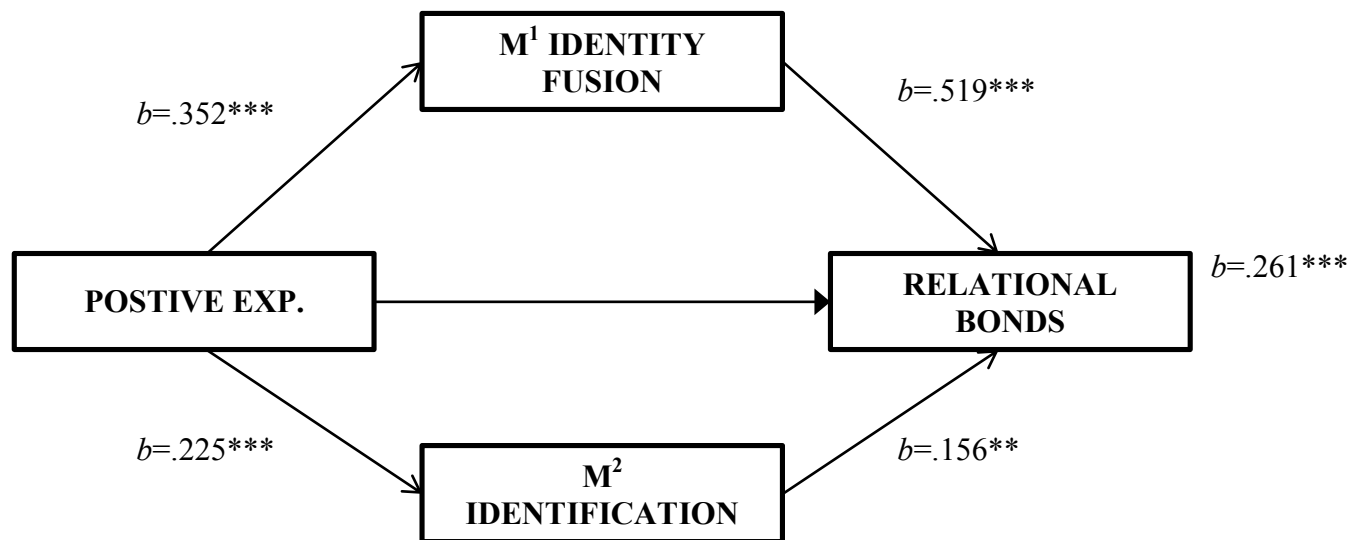
Table 2-23: Regression Models predicting relational bonds

Variable	Outcome: Relational Bonds					
	STAGE 1		STAGE 2		STAGE 3	
	B (SE)	β	B (SE)	β	B (SE)	β
Constant	1.329 (.340)	-	.661 (.365)	-	.429 (.335)	-
Train Hrs.	.054 (.011)	.222***	.047 (.010)	.192*	.025 (.009)	.101**
Promotion Attn.	-.014 (.006)	-.098*	-.011 (.006)	-.072	-.003 (.006)	-.019
Hor-Col.	.322 (.069)	.209***	.304 (.068)	.197***	.102 (.062)	.066
Ver-Col.	.194 (.051)	.168***	.151 (.051)	.131**	.070 (.046)	.061
Positive	-	-	.171 (.042)	.176***	.038 (.039)	.039
Negative	-	-	.065 (.039)	.070	.048 (.035)	.052
Identification	-	-	-	-	.142 (.053)	.121**
Fusion	-	-	-	-	.442 (.050)	.429***
R^2	.167		.201		.382	
F	23.474***		19.673***		36.054***	

$N=474$, $p<.05^*$, $p<.01^{**}$, $p<.001^{***}$ shading added to emphasis variables with significant coefficients.

The model that emerged in the second stage, including positive and negative measures along with other confounding variables, significantly predicted *relational bonds*; $F(6,468)=19.673, p<.001, R^2=.201$, with positive experiences making a significant contribution to the model ($\beta=.176, p<.001$). This represented a significant improvement over the initial model but the increase in explanatory power was minor; $\Delta R^2=.035, F(2,468)=10.229, p<.001$. But when group affiliation measures were added in the third stage the predictive power of the model increased substantially, $\Delta R^2=.181, F(2,466)=68.236, p<.001$. Interestingly, in this model the contribution of positive experience was superseded, with identity fusion becoming the strongest contributing variable (Table 2.23).

Figure 2-14: Fusion and identification mediate positive exp. and relational bonds



Direct effect $b=.044 (.038), p=.249$

Indirect effect M¹ $b=.183 (.028), 95\% \text{BCa CI } [.128, .242]$

Indirect effect M² $b=.035 (.015), 95\% \text{BCa CI } [.010, .073]$

Index of Mediation M¹ $\beta=.186 (.027), 95\% \text{BCa CI } [.134, .240]$

Index of Mediation M² $\beta=.036 (.016), 95\% \text{BCa CI } [.011, .074]$

$N=511$

Bootstrapping for Bias Corrected CI conducted with 1,000 Samples

To examine this potential mediating relationship a mediation analysis was conducted with identity fusion and group identification entered as two parallel mediating variables between the positive experience and *relational bonds* measures using PROCESS procedures (Hayes, 2012). A full mediation model emerged with significant indirect effects through both pathways. However, coefficients and standardized measures of effect size indicated that the mediation pathway through identity fusion was of a much greater magnitude.

In summary, counter to predictions, negative experience ratings did not predict *relational bonds* with other members of the BJJ school. Instead, positive promotion experience scores displayed a robust relationship with the *relational bonds* measure, and this connection was found to be fully mediated by identity fusion and identification, with identity fusion representing a much more robust pathway.

Hypothesis 6 (More Dysphoric = Higher Status & Solidarity)

To examine whether those who rated their promotion experiences as more negative also rated their groups as having a higher *status* (based on two independent single item measures) and greater solidarity (using a combined *entitativity* measure) in accordance with cognitive dissonance theories (Aronson & Mills, 1959; Festinger, 1962; Kamau, 2013), a correlation analysis was conducted between the two independent measures of perceived *group status* and the combined *entitativity* measure. Due to non-parametric distributions in the negative and positive ratings spearman's rho correlations are reported (Table 2.24).

Table 2-24: Spearman's correlations for group status and entitativity measures

	Status 1: Admired	Status 2: Respected	Entitativity
Negative Experience	.031	.010	.037
Positive Experience	.295***	.194***	.143**

$p < .05^*$, $p < .01^{**}$, $p < .001^{***}$

There was no correlation found between the negative experience scores and either of the *group status* measures or the entitativity measure. Instead, again counter to predictions, there was a weak correlation for all three measures with the positive experience rating. To examine whether this remained robust when controlling for other relevant factors a multiple regression was conducted on each of the three outcomes with the positive and negative experience measures entered as predictors along with demographic factors- age and sex; the binary measure to indicate the presence/absence of belt whipping; idiocentric/allocentric personality scores; group identification; average hours training per week; total years training; months since last personal promotion, and months since last grading attended.

Assumptions for all models were assessed prior to analysis: each model displayed independence of residuals and partial regression plots indicated approximate linear relationships with all non-categorical variables. There were also no indications of multicollinearity as assessed through correlation analysis and VIF and tolerance statistics. However, examinations of standardized residuals indicated 9 cases as extreme outliers with values over 3 SDs and a further 8 cases demonstrated extreme values of leverage and Mahalanobis distance. These 17 cases were removed for all models and analyses were rerun. Finally, visual examination of residuals through histograms, P-P plots and Normal Q-Q plots indicated a normal distribution for the *entitativity* and the '*admired*' *group status* measure. However, there was a severe positive skew with the '*respected*' *group status* measure. A square root transformation was applied to the scores but the distribution remained notably positively skewed and hence reported coefficients should be treated with caution.

Table 2-25: Regression models predicting group status and entitativity

Variable	Outcome:					
	Entitativity		Status 'Admired'		Status 'Respected' ⁴⁰	
	B (SE)	β	B (SE)	β	B (SE)	β
Constant	1.329 (.340)	-	.982 (.562)	-	2.265 (.195)	-
Sex	.123 (.249)	.023	.021 (.211)	.005	.098 (.073)	.064
Age	-.006 (.008)	.036	.020 (.007)	.146**	.003 (.002)	.063
Train Time Yrs.	-.039 (.017)	-.118*	-.022 (.015)	-.075	-.001 (.005)	-.012
Train Hrs.	.004 (.013)	.015	.030 (.011)	.124**	.009 (.004)	.117*
Belt Whip	-.079 (.107)	-.036	-.027 (.091)	-.014	.023 (.031)	.037
Last Attn.	.009 (.009)	-.050	.004 (.091)	-.050	.000 (.003)	.001
Last Promo.	.001 (.004)	.012	.002 (.004)	.036	.002 (.001)	.106*
Hor-Idv.	-.013 (.070)	-.009	-.019 (.059)	.015	.038 (.020)	.091
Ver-Idv.	.185 (.062)	.150**	.214 (.052)	.198***	.016 (.018)	.044
Hor-Col.	.159 (.076)	.102*	.107 (.064)	.078	.003 (.022)	.007
Ver-Col.	.132 (.058)	.116*	.050 (.049)	.050	.029 (.017)	.088
Positive	.114 (.046)	.120*	.189 (.039)	.228***	.049 (.013)	.178***
Negative	.037 (.045)	.041	.028 (.038)	.034	.005 (.013)	.018
R^2	.101		.161		.091	
F	3.684***		6.274***		3.266***	

$N=437$, $p<.05^*$, $p<.01^{**}$, $p<.001^{***}$ shading added to emphasis variables with significant coefficients.

The analyses reported in Table 2.25 indicated that when controlling for a variety of other potentially confounding factors, positive experiences remained predictive of both *group status* measures and the combined *entitativity* measure. This strongly increased confidence in the validity of the finding. Conversely, negative experience displayed no predictive power in any of the regression models. Although this finding is counter to the hypothesis provided, an important point of caution is that due to the nature of the questionnaire the positive and negative experience measures were asking respondents to rate a previously completed promotion experience. This is relevant as it is possible that dissonance effects could impact the assessment of the event, for instance, by leading them to generate more positive evaluations of the experience now that they have been promoted in to a more

⁴⁰ As this measure was reverse coded values were mirrored prior to analysis.

senior member of the group. This is especially relevant given that the constructed positive experience measure included items assessing the *value* and *meaningfulness* of the promotion experience.

This interpretation cannot be ruled out based on the current data but one point against it is that cognitive dissonance accounts have generally focused on individuals' responses to harsh group initiation events and have shown that it is typically the judgments of the relevant group, rather than the initiation experiences, which are adjusted.

Hypothesis 7 (Dysphoric promotion experiences = Greater reflexivity)

As part of the questionnaire respondents were asked to write an open ended response to their 'intuitions about the meaning of the BJJ belt promotion/grading ceremonies' providing details of their thoughts 'about the overall meaning of the grading ceremonies and/or the specific elements mentioned'. The responses were then rated by two hypothesis-blind, independent coders, on i) the amount of introspective reflection displayed and the extent to which the account provided indicated that belt promotions were described as ii) important, iii) traditional, and iv) pragmatic. These measures were selected in order to collect an indication of spontaneous exegetical reflection (SER) a crucial component of Modes theory (intended to be captured by the introspective reflection item) and three additional variables to enable coders to classify other aspects of the response and to allow other potentially confounding response factors to be considered. In particular, the introspective measure was intended as a direct measure of SER and coders were instructed to base their rating on how far they felt that the response demonstrated the respondent had thought personally/reflected on their experience. Coders were told not to prioritise length

but rather the depth, including analytical or metaphorical aspects, of responses. For the rating of importance coders were simply instructed to rate the degree of personal importance that the respondent appeared to attach to the experience. The instruction for rating traditional was to provide ratings on the extent to which the respondents made reference explicitly to BJJ traditions and history when discussing their promotion. This measure was anticipated to be orthogonal to the pragmatic rating, which was based on the extent to which respondents described their promotion event in instrumental terms (e.g. as an indication of skill progression/seniority, to provide structure to learning). However, the measures were recorded separately to avoid imposing an orthogonal relationship. The traditional and pragmatic measures were collected to help to distinguish between ritualistic, symbolic interpretations and more straightforward instrumental views. All four measures were coded using 6-point scales with the higher values indicating respectively more reflexivity and greater attribution of importance, tradition or pragmatism.

Table 2-26: Reliability of ratings

Variable	<i>r</i>	ICC (2,2) [95% CI]	α
Reflexivity	.734***	.824 [.701, .891]	.846
Importance	.668***	.802 [.706, .866]	.800
Traditional	.537***	.697 [.551, .796]	.696
Pragmatic	.363***	.415 [.399, .589]	.512

N=435, $p < .05^*$, $p < .01^{**}$, $p < .001^{***}$

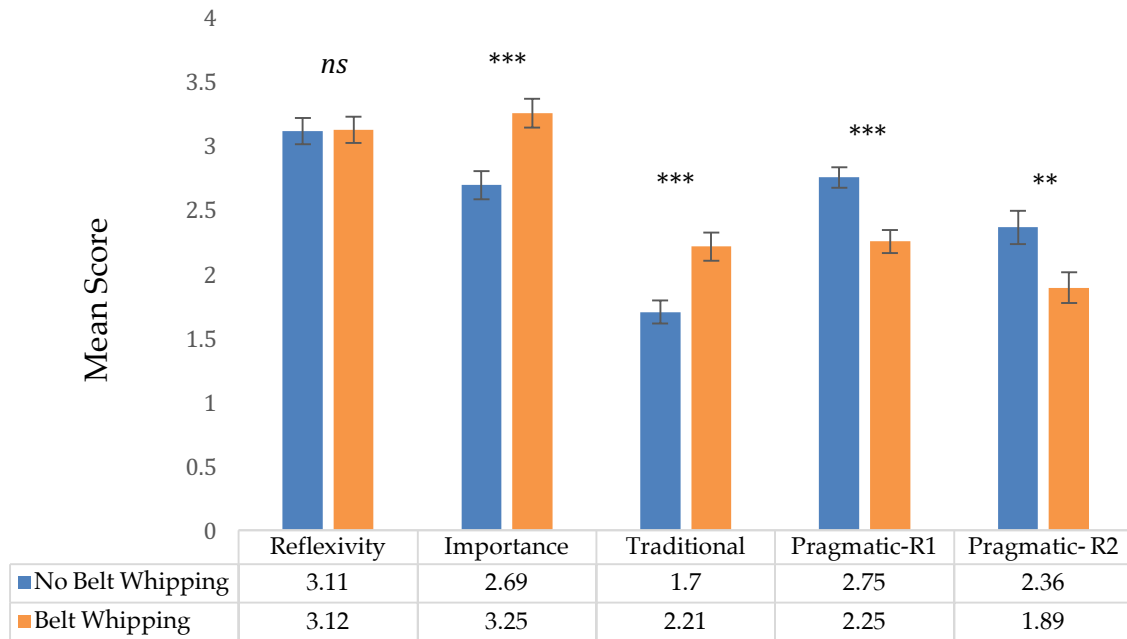
First, comparing results from the two raters, correlations were moderate for reflexivity and importance ratings and poor for traditional and pragmatic ratings (Table 2.26). Inter-rater agreement ratings (Landers, 2015), calculated using intraclass correlations (Shrout & Fleiss, 1979), indicated a high level of agreement for reflexivity and importance ratings, an acceptable level for traditional ratings and low levels of agreement for pragmatic ratings

(Table 2.26). Consequently, the reflexivity, importance, and traditional ratings supplied by the two coders were able to be combined to form a single rating score for each variable, while the pragmatism scores were treated independently for each coder.

Comparing the accounts provided by individuals with belt whipping promotion experiences against those with no whipping experience, those with belt whipping experience described their experiences as significantly more important, $t(435)=3.485$, $p<.001$, $d=.334$; more traditional, $t(435)=3.313$, $p<.001$, $d=.317$; and less pragmatic, Rater 1- $t(435)=-3.951$, $p<.001$, $d=.379$, Rater 2- $t(435)=-2.645$, $p=.008$, $d=.254$. The combined measure of reflexivity however was not found to differ significantly between groups; $t(435)=.061$, $p=.952$ (see Fig 2.15 for a summary).

The reflexivity results are counter to the findings reported in Richert et al. (2005) but this could be due to differences in the measures used between studies. The Richert et al. (2005) study specifically examined analogical depth of response and this was not assessed with the same level of fidelity in the current rating scheme. Moreover, the ratings in the current study were specifically addressing the reflexivity demonstrated in the description provided, this means that the importance measure may actually serve as a better proxy for broader reflexivity, since more important experiences may invite greater personal reflection, rather than more reflexive written descriptions. Some support for this is provided by the moderate correlation observed between the reflexivity and importance scores; $r=.577$, $p<.001$.

Figure 2-15: Ratings of reflexive accounts of promotional experiences



$N=435, p<.05^*, p<.01^{**}, p<.001^{***}$

Note: Error bars represent standard errors.

The difference in tradition ratings suggest that belt promotions that involve belt whippings are viewed as being more traditional. This may be an accurate assessment or it could be due to the ritualistic aspects and purported ancestry of the belt whipping practices. Supporting this interpretation both raters, despite the discrepancy in their ratings overall, found that individuals who did not experience belt whippings described their experiences in more pragmatic and hence instrumental terms. However, pragmatic and traditional ratings were not found to have an orthogonal relationship; Rater 1- $r=.153, p<.001$, Rater 2- $r=-.079, p=.101$.

To explore these findings a post-hoc hierarchical regression analysis was conducted to examine whether the negative experience score proved to be predictive of the *importance* attributed to accounts when controlling for a variety of relevant confounds. In the first

stage, potential confounds were entered, including demographic factors- sex and age; allocentric and idiocentric personality scores; training factors- time training in years, average hours training per week and salience factors- months since last grading attended and months since last promotion. In the second stage, the positive and negative experience measures were added and finally, in the third stage, the binary measure of belt whipping was added in order to assess whether the negative experience measure contributed to the predictive model independently from belt whipping.

Assumptions were tested prior to analysis. There was independence of residuals and partial regression plots indicated approximate linear relationships with all non-categorical variables, except for *months since last grading attended* and *months since last promotion*. However, exploratory analysis indicated that these variables did not significantly contribute to the model and they were thus not included in the final model. From the remaining variables there was no indication of multicollinearity, as assessed through correlation analysis and VIF and tolerance statistics. Examinations of standardized residuals and values of leverage and Mahalanobis distance also indicated no extreme outliers. Finally, visual examination of residuals in histograms, P-P plots, and Normal Q-Q plots indicated a normal distribution.

Results from the models (Table 2.27) indicated that only the vertical-individualist (VERIDV) idiocentric score was a significant confounding predictor of importance rating. Consequently, the variance accounted for in the model was small, $R^2=.042$. Adding positive and negative experience measures in the second stage substantially improved the model, and both were found to be significant predictors of the *importance* rated in accounts;

$F(2,374)=15.031$, $p<.001$, $\Delta R^2=.071$. The magnitude of the effect was slightly larger for the positive rather than negative experience ratings (Table 2.27).

Table 2-27: Hierarchical multiple regression on importance rating

Variable	Outcome: Rating of importance from subjective account					
	STAGE 1		STAGE 2		STAGE 3	
	B (SE)	β	B (SE)	β	B (SE)	β
Constant	1.559 (1.006)	-	-.238 (1.024)	-	-.881 (1.042)	-
Sex	.720 (.419)	.089	.769 (.406)	.095	.701 (.403)	.086
Age	-.016 (.013)	-.070	-.016 (.012)	-.067	-.014 (.012)	-.058
Train Time Yrs.	-.026 (.028)	-.051	-.025 (.027)	-.050	-.025 (.027)	-.049
Train Hrs.	.013 (.019)	.036	.048 (.010)	.001	.004 (.018)	.011
Hor-Idv.	-.056 (.115)	-.026	.315 (.052)	.011	.046 (.111)	.021
Ver-Idv.	.232 (.101)	.125*	.172 (.098)	.093	.161 (.098)	.087
Hor-Col.	.000 (.127)	.000	-.008 (.123)	-.003	.007 (.122)	.003
Ver-Col.	.157 (.095)	.090	.083 (.093)	.048	.070 (.092)	.040
Positive	-	-	.319 (.073)	.221***	.316 (.072)	.219***
Negative	-	-	.212 (.069)	.153**	.146 (.073)	.105*
BWhip	-	-	-	-	.478 (.173)	.143**
R^2	.042		.113		.131	
F	2.061***		4.778***		5.113***	

$N=384$, *= $p<.05$, **= $p<.01$, *** = $p<.001$, shading added to emphasis variables with significant coefficients.

Adding the binary belt whipping measure in the third stage only marginally improved the predictive power of the model, $\Delta R^2=.018$. But it was of note that in this model the negative measure continued to contribute significantly as this implied that negative experiences predicted *importance* ratings beyond the difference caused through the presence or absence of belt whipping practices. The positive experience rating displayed the strongest correlation and this remained robust in both models. This did not contradict expectations as negative and positive ratings did not display inverse correlations.

In summary, belt whipping promotions were not found to promote more reflexive written accounts but they did result in individuals making more references to *tradition* and

the *importance* of the promotion events. Non belt-whipping promotions were also described in significantly more *instrumental* terms. Further, post-hoc analysis also revealed that the subjective importance attached to promotions was partly predicted by both positive and negative experience ratings.

2.5 Discussion

The results from this sample demonstrated the anticipated variation in promotion practices, with 53% reporting that belt whipping gauntlets featured, 33.7% reporting being thrown, and 34.9% involving continuous sparring. This meant that the sample included an adequate amount of variation to enable examination of the stated hypotheses. The initial motivation for the study rested on the anticipation that the dysphoric and ritualistic nature of belt whipping promotional experiences would result in individuals who had endured such experiences displaying outcomes associated with the imagistic mode of religiosity (Whitehouse & Lanman, 2014).

Accordingly, an important initial step was to establish that belt whipping events were more dysphoric than other promotion events. This proved to be the case as the negative experience measure was found to be significantly higher for the belt-whipping group. Furthermore, there was no significant difference observed between individuals with belt whipping experience on the positive experience measure. This implies that while individuals did recognise the dysphoric aspects of belt whipping gauntlets, this recognition did not result in an overall less positive assessment of the event.

This interpretation is consistent with qualitative responses concerning the meaning that individuals perceived in their belt promotion practices. Content rating of these responses

demonstrated that on average those who endured belt whippings considered their promotions to be significantly more important and more traditional than those who did not have belt whipping experiences. However, such findings should not be interpreted as implying that there is universal support for the practice amongst those who have experienced belt whipping. On the contrary, the subjective responses recorded demonstrated a broad range of positions. To illustrate, consider the variation inherent in these three anonymous responses:

Respondent 1:

I believe that certain traditions and practices separate Brazilian Jiu-jitsu from other martial arts, particularly those watered down and or commercialized. BJJ is a truly special art that demands a special level of commitment; it is a unique and trying path...This is why the belt testing, and gauntlet tradition is so important. It is a reminder of our martial roots, and the tradition of perseverance BJJ demands. Not only does a whipping gauntlet turn away the weak of heart, it also demonstrates the difficulty of the path to come for the lower belts, and elevates those who have achieved a rank in their eyes. No one ever forgets their gauntlets, and I think for true practitioners, it is a source of pride.

Respondent 2:

Grading ceremonies for non black belts are generally meaningless... belt whipping is ridiculous, pointless garbage that should be confined to frat houses. Belts are meant to keep your gi top shut, and indicate roughly how you developed your jiu jitsu skill set is, not for obnoxious hazing.

Respondent 3:

I fail to understand the need to inflict pain or humiliation as a part of these ceremonies or those involved in BJJ. It certainly appeals to a segment of the population, but I am not one of them. I

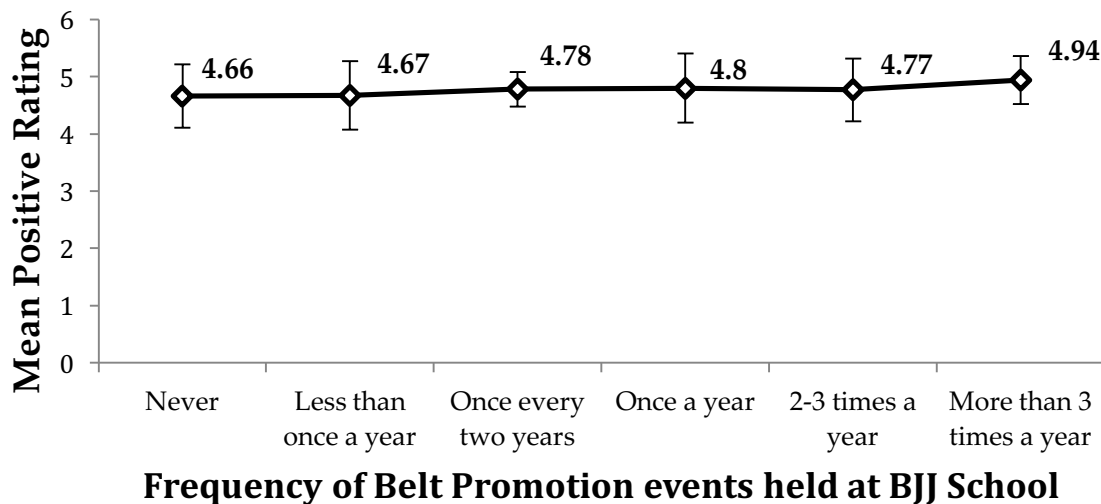
suffered intense physical abuse as a child and it's quite traumatic to endure these ceremonies. That is the primary reason I do not participate in whipping other students.

Overall, the most consistent finding across all analyses was that subjectively positively rated promotion experiences proved to be more strongly predictive than negative experiences for almost all outcomes of interest. This finding was largely unanticipated by the existing research literature and thus highlights an area that is in need of further investigation and potentially revisions to theoretical models.

There was however one variable on which the negative experience did display the predicted relationship: the frequency of belt promotion events. The hypothesis predicting this (H3) was derived from Modes theory; in particular, the results reported in Atkinson & Whitehouse (2011:56) that dysphoric arousal demonstrated a significant inverse relationship with ritual frequency, while euphoric arousal displayed a curvilinear relationship in a worldwide sample of rituals. This inverse relationship between frequency and levels of negatively valenced affect was replicated (Fig. 2.4) in the current data. Also in accordance with this hypothesis the positive experience ratings were not found to change significantly regardless of frequency of promotions (Fig. 2.16). The lack of variation in the positive experience measure is somewhat surprising, as it is reasonable to anticipate that promotion events that occur only once every few years would be considered more important events and generate a greater emotional response than those that reoccur every few months. But it could be that even the most frequent belt promotion events are still rare enough to generate significant levels of positivity. Ultimately all belt promotion events are celebratory occasions intended to recognise achievement and it may therefore simply be

that there is less scope for variance on the positive aspects. Even for those that hold critical views of the procedures of belt promotion procedures, as per the examples cited, a promotion event still provides a public signal of recognition of ability and years spent training.

Figure 2-16: Positive ratings across all frequencies of belt promotions



Note: Error bars represent standard deviations.

Whereas the relationship between dysphoric experiences and frequency accorded with predictions, the other connections hypothesised by Modes theory (Whitehouse & Lanman, 2014) between dysphoria and identity fusion (H1a, H2), relational bonds (H1b, H5b), and willingness to engage in personally costly sacrifices for the group (H5a) were generally not supported. Specifically, *relational bonds* and levels of fusion were not found to be significantly greater amongst individuals who experienced belt whipping, on both the verbal and pictorial measure of fusion (H1). Furthermore, in a finer grained analysis examining negative experiences across the full sample, no statistically significant correlation was observed with identity fusion. Indeed, in a hierarchical regression analysis

negative experience scores were found to interact negatively with the presence of belt whipping in promotions to predict lower fusion scores (H2). For *relational bonds* there was an initial weak correlation observed with negative experience ratings but further analysis controlling for relevant confounding factors found that this relationship was not statistically significant (H5b).

Focusing on the three personally costly pro-group sacrifice measures; initially a weak correlation between negative experiences and self-reported willingness to risk life for the BJJ school was observed ($r_s=.111$, $p=.011$). But this relationship was also not found to persist when controlling for relevant confounding factors (H5a). Comparing open ended responses (H7) to promotion experiences, although belt whipping experiences were not found to promote greater more *reflective* written accounts they did result in more references in the accounts being made to *tradition* and the *importance* of the promotion events. Non-belt-whipping promotions conversely were described in significantly more pragmatic/instrumental terms. These findings partly accord with previous results showing that individuals who had stronger affective responses to a dysphoric artificial ritual subsequently demonstrated greater depth of reflection on the experience (Richert et al., 2005). However, differences in the content coding employed make any reliable comparison impossible. In particular, the Richert et al. (2005) study coded responses from recorded oral interviews and measured specific linguistic aspects such as analogical depth, whereas the content analysis in this study involved shorter written responses and was based on codings that did not require specific linguistic expertise.

Focusing on the specific dysphoric element of pain, previous studies have shown that pain shared socially can enhance cooperation (Bastian, Jetten, & Ferris, 2014), perceptions of pain were associated with increased prosocial donations during a Hindu piercing ritual in Mauritius (Fischer & Xygalatas, 2014), and that anticipated pain increased generosity in imagined scenarios (Olivola & Shafir 2013). Therefore, it was expected that the pain of grading experiences would correlate with the three pro-group sacrifice measures, however no direct relationship was observed. This could be due to the lack of severity involved with promotion events when compared to more extreme events like piercing rituals, however this explanation would not account for the discrepancy with findings from the vignettes in Olivola & Shafir (2013) which used tiring physical activities to elicit expectations of pain. Another plausible explanation is that since the questionnaire was retrospective and the prosocial measures were being taken long after the events themselves had been completed, that dissonance reduction processes could have resulted in individuals reinterpreting their experiences in a more positive manner. Relatedly, in the three other studies referenced the relevant painful events had either just taken part or were being actively imagined rather than recalled and this may have resulted in greater saliency being attributed to the pain experienced (or imagined). Due to these differences and the possibility of dissonance reduction distorting accurate recall it would be premature to reject an association between pain and enhanced prosociality on the basis of these null findings.

Overall, in contrast to the lack of correlations with the negative measures, the positive experience measure displayed significant correlations with identity fusion (H2); including both visual; $r_s=.345$, $p<.001$, and verbal measures; $r_s=.357$, $p<.001$, *relational bonds* (H5b);

$r_s=.252$, $p<.001$, and all three costly sacrifice measures (H5a): Donate time- $r_s=.345$, $p<.001$; Risk Life- $r_s=.357$, $p<.001$, Donate money- $r_s=.252$, $p<.001$. Additionally, the relationship observed between positive experiences and the fusion measures proved to be robust when controlling for group identification, whereas conversely the correlation with group identification entirely disappeared when controlling for fusion. Confidence in the validity of the finding was also increased through a sub-group analysis of the belt whipping and non-belt whipping samples which detected a consistent relationship in both samples. Collectively, these results provide strong support for a relationship between identity fusion and positive promotional experiences and suggest that this connection was more substantial than that detected between positive experiences and categorical group identification.

In terms of the pro-group sacrifice measures, individuals with more positive promotion experiences displayed a greater willingness to endorse donating time, potentially risking their life for their BJJ school, and voluntarily donated more potential prize money (H5a). Although the reward offered in the study was relatively small (£20), participants spent on average 25-30 minutes completing the survey and thus invested a considerable amount of time with no immediate reward. In such a context, completing the additional (optional) questions to take part in the donation task, and then opting to donate some or all of the potential award is likely to reflect a genuine desire to donate to the group.

Significantly, initial regression analysis suggested a possible mediating effect of fusion between the positive experience ratings and the sacrifice measures. This was confirmed for both self-report sacrifice measures using mediation analyses which indicated that identity

fusion partially mediated the correlation with willingness to donate time and to risk life for the BJJ school. The significance of the mediation observed was also enhanced by the models used simultaneously controlling for mediation via categorical group identification. Categorical group identification displayed a mediating relationship between positive experiences and the willingness to donate time measure but this was weaker than the mediation pathway for identity fusion. Additionally, no significant mediating pathway for group identification was found for the more severe hypothetical sacrifice measure of willingness to risk life for the group (H5a). Finally, the bimodal distribution of the behavioural donation measure meant that a similar mediation analysis was not appropriate but consecutive ordinal regression models suggested a confounding relationship between fusion and positive experience ratings on the donation outcome.

Positive promotion experiences were also found to be weakly correlated with the constructed *relational bonds* measure; $r_s=.257$, $p<.001$, and this relationship was found to persist when controlling for a variety of other confounding factors in regression analysis. Again, follow up mediation analyses demonstrated an important mediating role for identity fusion; with the connection between positive experiences and *relational bonds* being fully mediated by identity fusion and identification, with identity fusion representing the more substantial mediation pathway.

These are highly significant results as they represent the first time that a clear relationship between *positive* experiences and identity fusion have been reported. They also offer initial evidence that identity fusion could function as a core mediating mechanism in translating positive ritualised experiences, and potentially euphoric arousal, into pro-group

outcomes and stronger relational bonds. This is a relationship that has heretofore not been emphasised in the identity fusion literature (Swann et al., 2014, 2009, 2012). Due to the cross-sectional nature of the data collected it is impossible to determine whether individuals experience more positive experiences at their belt promotions as a *consequence* of being more fused with their group and having stronger relational bonds or alternatively whether positive experiences at belt promotions served to generate greater levels of identity fusion. To clarify this the causal relationships will require further examination through controlled experiments.

Finding a relationship between positive affect and a measure of social bonding, however, does accord with the broader literature on affect and social bonding, briefly reviewed in section 1.8, which has collated substantial evidence that positive emotions serve to enhance prosociality and facilitate social bonding (Isen, 2000; Kelly et al., 2014; Parkinson, Fischer, & Manstead, 2005; Shiota et al., 2004; Spoor & Kelly, 2004). Furthermore, collective promotional events seem to provide an ideal environment for the occurrence of *emotional contagion* (Barsade, 2002; Hatfield, Bensman, Thornton, & Rapson, 2014; Hatfield et al., 1994), which refers to a theorised tendency to automatically mimic others expressions, vocalisations, postures and behaviours and thus to converge emotionally. In a sporting context previous research has shown, for example, that the average of teammates' happy moods are linked with individuals' own moods and subjective performances (Totterdell, 2000). Thus, it is possible that the celebratory nature of receiving a promotion results in positive affect being transmitted amongst BJJ practitioners during promotion events. Additionally, emotional contagion processes apply not only to

positive but negative affect and this implies that displays of exertion, pain, and discomfort at BJJ promotions should also have an impact on individuals who observe them. The implication then is that BJJ promotion events overall represent more positive affecting experiences and that this encourages the formation of stronger relationships and more positive associations with other group members present.

However, one potential confounding factor is that the vast majority of BJJ practitioners intrinsically recognise promotion events as positive experiences and thus it may be that even if they object to aspects of the event, such as the belt whipping gauntlets, or have a negative personal experience, practitioners will readily report promotion events as positive. In support of this, based on the single item self-report measures 62% ($N=354$) of respondents rated promotion events as being very valuable or extremely valuable, whereas only 12.1% ($N=73$) rated them as not valuable at all or a little valuable. Similarly, 72% ($N=424$) rated them as very meaningful or extremely meaningful, versus 7.1% ($N=43$) who rated them as not meaningful at all or a little meaningful and 68.4% ($N=414$) rated them as very enjoyable or extremely enjoyable, versus 7.2% ($N=42$) who rated them as not enjoyable at all or a little enjoyable. It would be interesting to know if endorsement of the practices and fusion with the BJJ school were correlated but endorsement was not specifically measured in this study.

Further evidence that experiencing dysphoric aspects of promotions did not necessarily equate to a decrease in positive experiences was supplied by the lack of an inverse correlation between positive and negative experience scores; $r=.072$, $p=.095$, and a weak inverse correlation between positive experiences and pain ratings; $r=.085$, $p=.049$.

Ultimately, BJJ is a combat sport with a Brazilian heritage, and there is a strong culture of machismo (Baldwin & DeSouza, 2001) that pervades the community. Within BJJ schools those who are capable of enduring pain and discomfort with stoicism or good humour are highly praised. Hence, although the questionnaire was conducted online and emphasised anonymity to respondents, it is still likely that the salience of group norms related to toughness and stoicism could reduce willingness to report belt promotion events as being painful or unpleasant.

In support of this interpretation, two recent studies conducted at firewalking events found that for performers “enduring a ritual ordeal was associated with greater happiness...post-ritual” (Fischer et al., 2014) and there was a “spectacular discrepancy between actual heart rates and self-reported arousal” (Xygalatas et al., 2013a). In the second study, fire-walkers in Greece consistently under reported their physiological arousal and Xygalatas et al. (2013a: 12) suggest that this trend was due to surrounding cultural norms imposing a “strong expressive suppression on participants”. This in turn is hypothesised to exhaust mental resources for forming accurate memories under the strain of exercising emotional regulation (Schjødt et al., 2013). From the qualitative accounts collected and the personal training experiences of the study author it is clear that similar framing exists in BJJ culture. Amongst BJJ schools that practice belt whipping gauntlets, or other similar dysphoric practices, there is a clear effort made by those who take part to appear resilient and unreactive while performing the gauntlets⁴¹.

⁴¹ There are some members who react dramatically, usually to humorous effect, and also a smaller number who will ask to be excluded from such events. However, the majority comply with expectations and even

It is possible therefore that the positive experience reported is actually the result of inaccurate recollections derived from efforts to down regulate emotional responses experienced when under physiological stress (Schjødt et al., 2013). There are some problems with this interpretation, however, as unlike the firewalkers described in Xygalatas et al. (2013a), after the event has taken place, recipients of belt-whippings are often quite willing to joke and complain about the pain involved. This is likely a consequence of the recreational nature of BJJ communities but it suggests that the suppression of emotion is less important with BJJ practitioners than it is with the firewalkers surveyed. Whether a willingness to comment on the pain of the experience in a training context when surrounded by fellow group members would translate into a willingness to report negative experiences in an anonymous survey is difficult to assess. In short, due to the potential overlapping nature of affective responses to dysphoric events, such as belt whippings, and the presence of a cultural frame that encourages stoicism and rewards displays of endurance, it would be advisable to avoid over-interpreting self-report of experiences as being accurate indicators of the actual severity of promotional practices.

Although caution is warranted it is equally important to be cautious of motivated reasoning conveniently explain away incongruent findings (Kunda, 1987). That is, if the same robust correlations had been observed between negative experiences and the anticipated outcome measures it is less likely that the significance of such findings would be dismissed due to the limitations of self-report methods. It is not the intention therefore

individuals who are derisive of belt whipping practices in discussions tend to behave according to group expectations when promotion events take place.

to seek to minimize the significance of the relationships observed between positive experience, identity fusion, and all three of the prosocial sacrifice measures. The results are preliminary, involving a potentially atypical community, and a ritual event without sacred connotations but it remains striking that counter to the current emphasis in the most recent versions of Modes theory dysphoric promotion experiences were not found to be predictive of identity fusion or greater willingness to make costly sacrifices for the group. Instead, it was positive promotion experiences that displayed these predicted relationships. The extent to which these findings then can be considered compatible with Modes theory will be discussed further in the conclusion.

Turning first to address identity fusion theory directly, the results obtained for H4 provided strong support for the core predictions of the theory; demonstrating that identity fusion measure was more strongly associated with all three sacrifice measures than the categorical identification measure. Individuals with higher levels of fusion displayed a greater willingness to donate time, risk their life, and voluntarily donate prize winnings to their BJJ school. Additionally, in accordance with existing studies demonstrating a mediating role of perceptions of familial ties (Buhrmester et al., 2014; Swann et al., 2014), relational bonds were found to act as partial mediators between fusion and both self-reported sacrifice measures.

Drawing on connections proposed recently by Whitehouse & Lanman (2014), subjective ratings of meaningfulness, intended to serve as a proxy for reflection on the experience, were also entered in the mediation models and a partial mediation was observed with the more extreme willingness to risk life measure. That there was no significant mediation with

the time donation measure suggests that the pathway is less robust than with relational bonds but it should also be noted that fusion, empirically and theoretically, is considered to be a stronger predictor of *extreme* pro-group behaviour. Therefore, it may be that the personal meaning and spontaneous reflection attached to belt promotions becomes more relevant for highly fused individuals specifically when they are being asked about more extreme personally costly pro-group behaviors.

When examining the hypothetical behavioural donation measure no mediation was detected for *relational bonds* or meaningfulness in an exploratory mediation analysis however the non-parametric distribution of the outcome variable makes these findings unreliable. Furthermore, although intended to serve as a costly measure the donation was ultimately a hypothetically donation and involved only a potential loss of a relatively low reward payment. Hence, even if the preliminary results found were valid they do not clearly relate to predictions based on extreme costly behaviours.

In short, the primary predictions of identity fusion theory were supported and preliminary evidence was provided for a potentially mediating effect for the meaningfulness attached to important group events. However, these findings need to be explored using more precise measures in independent samples before any broader conclusions can be drawn.

Drawing on research from cognitive dissonance, automatic accrual, and costly signalling theories it was hypothesised (H6) that respondents would assess their BJJ schools as having higher status and greater solidarity if they had participated in more dysphoric promotion experiences. Cognitive dissonance theory proposes that, to reduce feelings of

dissonance, those who perform painful or effortful activities for their group should evaluate their group more positively (Aronson & Mills, 1959; Gerard & Mathewson, 1966; Harmon-Jones & Harmon-Jones, 2007). Alternatively, one costly signalling interpretation (Bulbulia & Sosis, 2011) could be that high status groups will require more costly demonstrations of commitment (CREDs) (Henrich, 2009) in order for veteran members to prevent non-committed newcomers from gaining access to valuable group resources, including associated prestige (Cimino, 2011, 2013). Even though the causal direction between these two accounts differ, the outcome predicted is the same; groups with more dysphoric belt promotions should be rated as having a higher status by those who belong to the group and potentially as offering more benefits, such as increased solidarity.

However, no significant correlations were observed with the negative experience measure and either the combined *entitativity* measure or the two *group status* measures. Instead, mirroring the pattern reported in the previous hypotheses, significant correlations were observed with the positive experience measure and all three outcomes. The strength of the correlations was relatively weak (*group status*- admiration, $r=.295$, $p<.001$; *group status*-respect $r=.194$, $p<.001$; *entitativity*, $r=.143$, $p<.001$) but this is not surprising given that there is a wide assortment of unrelated factors that influence a BJJ school's relative status, such as an instructor's achievements, success in competitions, size of the group, quality of facilities, etc. Furthermore, confidence in this relationship was increased through follow up regression analyses which demonstrated that positive experiences remained substantial predictive factors when controlling for a variety of demographic, personality, and experience measures. In summary, this finding emphasises the importance of shared

positive experiences in promoting positive images of the relevant group, a non-controversial suggestion, but one that has been somewhat overlooked in recent research on high-arousal ritual events.

A possible alternative interpretation is that the causal arrow is reversed with higher status groups tending to have more positive belt promotion events due to the controversial nature of belt whipping and other dysphoric promotional activities. There is some preliminary evidence in support of this as there was a significant correlation between schools with larger numbers of participants in promotion events and the *group status-admiration* measure; $r=.168, p=.004$. However, no such correlation was found with the *group status-respect* measure; $r=-.084, p=.156$. Moreover, school size is not necessarily a suitable proxy to indicate that a school holds a high status within the BJJ community. Larger clubs can be accused of promoting personality cults focused on a small cadre of high-level professional competitors (Rossen, 2013) and of failing to deliver personalised instruction (Somnez, 2009). Again, further data is needed before any reliable conclusions can be drawn.

Additionally, it is generally not possible for a student to assess a school's quality on the basis of how harsh the promotional practices are since these are not advertised and most newcomers are unaware of what promotion entails when they join a school. In the current sample 62.7% ($N=379$) reported having no idea about their school's promotional practices before joining, while a further 16.1% ($N=97$) had only a vague idea. Only 6.9% ($N=42$) indicated that they knew anything about their school's specific promotion events. This makes any signalling account that would identify harsher promotion events as a means for

providing a public signal to encourage the recruitment of high quality, devoted members as problematic.

Considering the belt whipping promotion practices as a potential hazing initiation event, it is evident that they meet almost all of the criteria listed by Cimino (2013). They are *temporary* events used to signal a change of status and while participation is theoretically voluntary it is often practically *coerced*. Furthermore, they are directed at deepening an affiliation with a long-term *coalition* that has typically lasted over generations. However, belt-whipping is not necessarily *uni-directional*, in that although it is considered bad form for new members to join in with too much enthusiasm, junior members are still often permitted to whip more senior members. Moreover, general group affiliation is not predicated on promotion events, since individuals can be training at a school for a substantial period of time before they experience any promotion event. In summary then it does seem appropriate to apply Cimino's model to understand the practice but a core issue is that the key motivation identified by this model appears to be missing. Specifically, individuals are able to gain access to the group and associated benefits *prior* to participating in any belt promotion event.

However, it could be that the more significant automatic benefits, such as associated status or prestige, only really begin to accrue after practitioners receive formal recognition (they are awarded their first belt) from their school. In support of this, it is common in the BJJ community to discuss the general level of blue belt practitioners (the first earned belt grade) that a school produces as being indicative of a school's overall quality. Thus, BJJ

schools and individual instructors have intergroup motivations to only promote students that will reflect positively on the school's standards in competitions.

Conclusion

The results discussed above are preliminary and are drawn from an atypical martial arts community but the initial inference from the findings is that the present emphasis on dysphoric ritual events as a precursor of fusion in Modes theory may not be warranted. Conversely, there is strong support in favour of examining potential connections between positive experiences and identity fusion and this by extension implies that connections between euphoric rituals and identity fusion should not be overlooked. A recent commentary by Xygalatas (2014) on Modes theory and different forms of bonding (Whitehouse & Lanman, 2014) echoed this sentiment, noting that there is presently a paucity of evidence concerning the effects of euphoric high arousal rituals and, hence, there is no reason to believe "that euphoric rituals could not have similar effects" as those anticipated for dysphoric events (Xygalatas, 2014: 16).

Despite the recent emphasis on shared dysphoric experiences in generating identity fusion (Swann et al., 2012; Whitehouse & Lanman, 2014), it may be that a broader connection exists between imagistic high-arousal rituals in general and identity fusion. This would accord with earlier formulations of Modes theory (Whitehouse, 1996, 2000) which placed an emphasis on the ability of high arousal imagistic rituals to generate rich, personally consequential episodic memories. The theory goes as follows: highly arousing imagistic rituals, which are typically not accompanied by official doctrinal interpretations, encourage spontaneous exegetical reflection (SER) (Richert et al., 2005) and this leads to

participants in forming rich self-defining autobiographical memories (Whitehouse & Lanman, 2014). Perceiving others to have shared in events that are viewed as self-defining in turn then leads to the development of social bonds based on a shared sense of psychological kinship which in turn can potentially lead to identity fusion.

In this account there is no need to posit that highly arousing positive ritual experiences are unsuitable for establishing the self-defining memories that are at the core of the proposed relationship with identity fusion. Moreover, as this chapter and recent findings with Mauritian firewalkers attest (Fischer et al., 2014), dysphoric ritual events are capable of generating subjective positive responses. This does not make the objective distinctions between dysphoric and euphoric rituals irrelevant but it does suggest that researchers should be careful to avoid labelling dysphoric rituals as if they necessarily entail the generation of internal negative affect. When walking through fire or being whipped with belts it is undeniable that the body is experiencing a physiologically stressful situation and that this will stimulate psychological and hormonal responses. However, it is equally clear that such responses can be strongly moderated by an individual's internally generated response to the stress. For certain people enduring pain and discomfort may actually be experienced as pleasurable or alternatively to generate a contrast that results in a post-event euphoria (Bastian, Jetten, Hornsey, et al., 2014). This should not be a controversial suggestion as it is attested to not only in ritual contexts but also by the popularity of various popular activities such as attending horror films, bungee jumping, and enduring marathons (Boecker et al., 2008; Kerr, 2015).

In summary then the core argument of this chapter is that there is no fundamental theoretical reason in the connections posited by Modes theory that require either that imagistic euphoric rituals are unable to operate through the same processes as those proposed for dysphoric imagistic rituals or that dysphoric rituals are not capable of generating pro-social behavior and social bonding outcomes through positive subjective responses. Yet despite this it is also evident that the connections between positive affect, fusion, and group orientated prosociality have not yet been addressed in sufficient depth in the existing research literature. Based on the findings of the present chapter, this seems like a productive avenue to pursue further.

This chapter examined a variety of predictions against data from promotional events in a novel martial arts context and provided strong evidence in support for the established relationship between identity fusion and prosocial sacrifice, using both self-report and behavioural measures (Gómez, Brooks, et al., 2011; Swann et al., 2014, 2009). Additionally, the meaningfulness of an experience and the *relational bonds* felt with other participants were found to mediate the observed connection between fusion and costly sacrifice for the group. These findings accorded with recent key predictions of Modes theory, namely that imagistic ritual events are effective at generating self-defining memories and that perceiving to share these with others invites psychological kinship operating through identity fusion.

The most striking result however was that, counter to predictions, dysphoric experiences were not found not to be predictive of identity fusion, relational bonds, or willingness to make costly sacrifices for the group. Instead, it was positive experiences that

displayed the predicted connections with these variables. Moreover, the connections between positive experiences and costly sacrifices were mediated in accordance with Modes theory by identity fusion rather than group identification. Predictions derived from dissonance and costly signalling accounts that more severe promotional events should be associated with higher evaluations of the relevant school were not found to be supported.

There are many potential alternative explanations for the results reported including the recreational nature of BJJ communities, the lack of sacred symbolism in promotion events, the confounding heterogeneity of the sample, and so on. Yet, Modes theory has sought to position itself as a theory that applies beyond religious traditions and BJJ belt whipping gauntlets include most of the characteristic features of imagistic dysphoric ritual events. Consequently, the findings reported remain relevant and suggest that there is a need for closer attention to the role, and potential affiliative generating effects, of positive experiences even when they occur in dysphoric high arousal contexts. Future hypotheses need to address this ambiguity by being clearer to specify whether hypothesised connections relate to subjective affective responses or are targeted at the more objective level of the presence of dysphoric activities.

CHAPTER 3: ARTIFICIAL RITUAL EXPERIMENTS

This chapter reports the results of two novel artificial ritual experimental studies that employed affective manipulations to examine the psychological effects of negative and positive arousal in a collective ritual context. The first experiment was conducted in the UK and was intended as a feasibility test of the artificial ritual paradigm and the second experiment was conducted in Japan and employed a revised methodology which included stronger stimuli, better controls, and new dependent measures. Initially these studies were intended to serve as a cross-cultural comparative study but due to methodological issues identified during the UK study there were significant changes in methodology introduced between sites. This meant that the two experiments were not examined as part of a single study but instead represented two independent studies connected by a core methodology that was refined in the Japan sample.

3.1 Background & Rationale

The previous chapter explored theories concerning the bonding effects of high arousal promotional events within martial arts communities and the results collected suggested that subjective positive assessments of promotional experiences, regardless of their dysphoric content, are important for the promotion of strong relational bonds, identity fusion, and prosocial ingroup behaviour. However, there were significant limitations with the data collected. First, the study was correlational making any causal interpretations questionable. Second, there was substantial heterogeneity in the sample due to the diversity of BJJ groups and the online sampling method employed. Third, the recreational nature of most martial arts communities, combined with the positive associations with belt

progression, may mean that the trends observed are specific to promotional events and do not generalise to a wider ritual context. The experiments reported in this chapter were intended to help address these limitations by examining whether the significance of positive subjective experiences was observed in a novel collective ritual, performed with strangers, under tightly controlled experimental conditions.

Although the experiments in this chapter build on the findings from the BJJ survey the primary motivation for the design was a desire to experimentally examine the effect of collective dysphoria in a ritual context on group bonding and prosociality. To elaborate, the experiments were intended to provide direct tests of recent predictions that dysphoric arousal in collective rituals results in a distinctive form of group bonding that is based on identity fusion and promotes extreme pro-group behaviour (Whitehouse & Lanman, 2014; Whitehouse & McQuinn, 2012). Modes theory has long argued that high arousal rituals can produce distinctive effects from their low arousal counterparts (Whitehouse, 1995, 2000, 2004) but more recent summaries suggested that dysphoric arousal is the key to generating the intense relational bonds and identity fusion (Whitehouse & Lanman, 2014; Whitehouse & McQuinn, 2012). These predictions are partly derived from Atkinson & Whitehouse's (2011) observation that the inverse correlation between high arousal and low frequency in their worldwide ritual sample was primarily driven by negatively valenced high arousal rituals. Furthermore, a number of researchers have proposed that there are important distinctions between the roles of euphoric and dysphoric arousal in religions and rituals (Y. I. Russell et al., 2011, 2014; Whitehouse & Lanman, 2014).

Unfortunately, there have currently only been a small number of studies that have attempted to explore these issues experimentally. To provide a short overview of relevant findings: Russell et al. (2014) found that differential affect promoted distinctive “cognitive response styles” to a ritual-like task. In particular, dysphoric affect was associated with stronger analogical reasoning, when it was paired with relevant expertise. These findings are preliminary and are based on a priming methodology which manipulated affect entirely independently from the proxy ritual task but they accord with previous findings from Richert et al. (2005) that high arousal ritual experiences promote greater analogical depth.

The Richert et al. (2005) study despite being conducted earlier also addressed some of the limitations with the Russell et al. (2014) study. In particular, it involved participants performing ritual procedures, during which aspects of the procedure and surrounding environmental features were manipulated to create a high (dysphoric) arousal and low arousal version. The study was exploring Modes theory’s predictions related to spontaneous self-reflection and found that participants who experienced stronger physiological arousal produced a greater volume and depth of analogical reflection during post experiment interviews and that these effects increased in magnitude after a delay. Due to the dysphoric nature of the high arousal version of the ritual, the results provide initial support that dysphoric events prompt stronger post-event reflection which, as discussed, has been hypothesised in Modes theory to be a key component in the forging of autobiographical memories and promoting strong relational bonds with other ritual participants. Methodologically the study also established that participation in one-off

artificial rituals, even when performed with strangers, could produce significant psychological and physiological effects.

As there was no euphoric condition in this experiment it is impossible to know whether the effects observed derived from general high arousal or are specific to dysphoric high arousal. The study also did not actually detect any difference in level of reflection between the low and high arousal experimental conditions. Instead, the significant results reported relied on post-hoc groupings based on the magnitude of physiological response experienced. Although this was a justifiable strategy, the results should be treated with caution unless supported by further experimental evidence or independent replications. The lack of difference between conditions also implied that stronger experimental manipulations than those reported were required to generate distinctive affective conditions.

In summary, the Richert et al. (2005) and Russell et al. (2014) studies both directly address the topic of dysphoric arousal in a ritual context and present some comparable results concerning dysphoric arousal's ability to promote analogical reasoning. But both studies also have significant limitations and neither explicitly addressed the topic of group cohesion⁴². Unfortunately, there have not yet been any published attempts to independently replicate the Richert et al. (2005) methodology but another recent study by Lang et al. (2015) has recently demonstrated that inducing anxiety increased ritualized behaviours in a spontaneous cleaning task. The study thus provides additional evidence for a connection

⁴² In the case of the Russell et al. (2014) study the experiment was performed individually, hence there was no group to bond with.

between negative affect and ritualised behaviour, however the causality is reversed and the Lang et al. propose that the link is based on pathways related to hazard avoidance (Boyer & Liénard, 2006; Liénard & Boyer, 2006), rather than the group bonding pathway presented by Modes theory. The results of the Lang et al. study serve as a potent reminder that there are likely to be multiple interacting pathways between negative affect, ritual performances, and ritualised behaviours and that care must be taken not to construct overly deterministic models when there is currently a paucity of evidence.

One recent study (Wen, Herrmann, & Legare, 2016) that did explicitly address ritual and bonding, involved groups of children performing a quasi-ritual task and then examined “the impact of ritual participation on children’s in-group affiliation”. The study compared groups of children in a day-care setting who participated in a necklace making task that was either “scripted, synchronous” and “demonstrated by a group leader” (ritual condition) or “non-scripted” and “supervised [but not demonstrated] by a group leader” (control condition) and found that participating in the ritual “increases in-group affiliation to a greater degree than group activity alone” (Wen et al., 2016). Meaning that children in the ritual condition displayed higher in-group identity fusion scores, made more affiliative choices, and had greater expectations of being included by their in-group. These results are consistent with connections posited in recent versions of Modes theory between ritual actions and identity fusion (Whitehouse & Lanman, 2014) however the methodology involved no affective manipulation and thus the key component hypothesised to be necessary for generating fusion was absent. There was also no measure of group

identification collected, so it is unclear whether the bonding observed was specific to identity fusion, or reflected a more general affiliative tendency.

Another potential limitation of the study is that according to the definition supplied by the authors ritual referred to “conventional, causally opaque procedures... uninterruptable from the perspective of physical causality because they lack an intuitive or observable causal connection between the specific action performed (e.g., rubbing a ceramic pot) and the desired outcome or effect (e.g., making it rain)” (Wen et al., 2016). The issue with this is that the intended *ritual* actions took place within a broader framework of children being asked to copy a leader’s instructions and ‘play’ with beads. Such framing could have potentially provided a coherent causal connection between the actions performed and a desired goal (to play, to follow a leader) and thus call into question to what extent causal opacity was generated. Finally, another possible alternative explanation for the effects observed is that they derive from the additional attention being demonstrated by an authority figure in the ritual condition. However, despite these possible limitations, the Wen et al. (2015) study still offers further evidence for the utility of a between-group experimental ritual methodology and some additional preliminary evidence for a connection between collective ritualized behaviour and affiliative tendencies.

Finally, examining actual ritual performances during the Hindu Thaipusam festival in Mauritius, Xygalatas et al.'s (2013b) experimental field research provided the first real-world evidence that performing or observing in extreme dysphoric rituals increases prosocial behaviour and inclusive self-identification, when compared to participation in low arousal rituals. A potentially analogous result that the collective sharing of pain acted

as a form of 'social glue' promoting "trusting interpersonal relationships by increase[ing] perceived bonding... [and] cooperation in an economic game" was also found in a recent lab-based study (Bastian, Jetten, & Ferris, 2014: 2080).

Collectively these experimental and field based findings, when considered alongside the trends reported in Atkinson & Whitehouse's (2011), offer a strong case justifying the examination of whether there is a distinctive role for collective, highly arousing *dysphoric* rituals in the formation of interpersonal bonds. The current evidence is broadly positive but remains preliminary and there have been no independent replications of most of the results discussed. Moreover, there has so far been no experimental study that has directly manipulated affect within a ritual context with the intention of examining group-cognition and affiliative behaviours despite Atkinson & Whitehouse (2011:60) highlighting the need for this direct kind of experimental test a number of years ago: "it should be possible to measure and directly manipulate subjects' arousal during simulated ritual participation and test how this is related to... ingroup vs. outgroup attitudes and cooperation".

The two studies reported in this chapter attempt to address this gap in the literature by presenting two between-group experimental studies that involved participants performing collective ritual sequences, while arousal was experimentally manipulated through a variety of environmental effects (Study 1: sound effects; Study 2: sound effects, lighting, and schematic faces). The primary goal of the studies was to isolate the impact of affect in a collective ritual and to test the connections proposed between dysphoric arousal and group bonding from the theoretical literature (Bulbulia & Sosis, 2011; Whitehouse & Lanman,

2014; Whitehouse, 2004) and recent field experiment findings (Fischer & Xygalatas, 2014; Xygalatas al., 2013b).

Accordingly, a variety of self-report and behavioural measures explicitly related to group cognition were collected, including; ingroup cooperation, group identification and identity fusion (Study 1), and intergroup discrimination and implicit positive associations with other (fictional) group members (Study 2). Methodologically these studies sought to build on the foundation of Richert et al.'s study (2005) and develop an experimental task that involved a collectively performed *ritual* that also camouflaged the affective manipulations as part of the ritual context through the use of a simple cover story.

Study 1: 'Wailing Infant' Artificial Ritual V1 (UK)

Study 1 involved a between groups experiment conducted in the UK, in which small groups of three participants performed a collective, artificial ritual task while surrounding levels of thematic affect (Czachesz, 2010) were manipulated through audio stimulus. Specifically, the study employed infant vocalisations to elicit differential affect: an unpleasant sound (a distressed, crying baby) for a dysphoric condition, and a neutral sound (a content, gurgling baby) for a control condition. A planned euphoric condition using a laughing baby sound was dropped during pre-testing of stimulus as responses to the euphoric stimuli were not found to differ significantly from the control condition based on the reported responses gathered from 22 hypothesis blind lab members. Additional details of the pre-testing of stimulus are provided in the methodology section. Overall, a priority in the design of this first study was to produce a simple, streamlined experiment that could be completed swiftly with few experimenters, required little sophisticated

equipment, and thus only contained the core relevant measures. Part of the motivation for this streamlined design was to enable the study to be easily adapted for other cultural contexts and also because it was envisioned to serve as a framework that would be employed in various follow up studies that added in additional measures of interest. Consequently, the audio track was the only feature which varied between the two conditions.

Infant vocalisations were chosen for the affective stimulus as an infant's cry is one of the most emotionally salient and aversive sounds in our environment (Boukydis & Burgess, 1982; A. D. Murray, 1979; D Out, Pieper, Bakermans-Kranenburg, & van Ijzendoorn, 2010; Philip Sanford Zeskind, 1987), possessing an unique potency to mobilize adult responses (Parsons, Young, Parsons, Stein, & Kringelbach, 2012; Zeifman, 2001; Philip S Zeskind, Klein, & Marshall, 1992; Philip Sanford Zeskind & Barr, 1997). Furthermore, aversive responses to infant crying sounds appear not to be culturally restricted (A. D. Murray, 1985; Philip Sanford Zeskind, 1983), and have been found to elicit the release of the hormone cortisol in men and women, regardless of parenting experience or age (Fleming, Corter, Stallings, & Steiner, 2002; Giardino, Gonzalez, Steiner, & Fleming, 2008)⁴³. Cortisol is a hormone usually associated with stress (Dickerson & Kemeny, 2004) but has also been found to be associated with more affectionate contact behaviors and more caretaking behaviors amongst first time mothers (Fleming, Steiner, & Corter, 1997), and greater

⁴³ This is not to minimize the existence of well documented differences in the magnitude of response between parents/non-parents (Lounsbury & Bates, 1982; Dorothee Out, Pieper, Bakermans-Kranenburg, Zeskind, & van Ijzendoorn, 2010; Seifritz et al., 2003), or in the potential for infant cries to elicit variously empathy, distress, and irritation, depending on personality and experience (Lin & McFatter, 2012; Dorothee Out et al., 2010). This was recognized as a potential confound and was controlled for.

sympathy in response to hunger cries among more experienced mothers (Stallings, Fleming, Corter, Worthman, & Steiner, 2001).

Following the ritual task, participants completed a short questionnaire about their experience, the meaning of the ritual, and participated in a five-round public goods game intended to operationalise levels of ingroup cooperation (following a three-day delay). The three-day delay was based on the theoretical claim of Modes theory that providing time for introspective reflection on a ritual event is central to the generation of interpersonal bonding effects (Richert et al., 2005; Whitehouse, 2004). Incorporating a delay was considered to be a truer test of the theory since the puzzling nature of the experience and the heightened arousal from performing tasks with a group of strangers may have restricted participants' abilities to spend time reflecting on their experience and solidifying an episodic memory. Heart rate measures were collected in order to provide a simple physiological indication of arousal through the use of wireless heart rate transmitters worn around the chest.

Hypotheses

H1: The negative/dysphoric condition will promote higher levels of a) identity fusion and b) perceptions of relational bonds with other participants (Whitehouse & Lanman, 2014).

H2: There will be greater cooperation between participants in the negative/dysphoric condition in the five round public goods game (Whitehouse & Lanman, 2014; Wiltermuth & Heath, 2009; Xygalatas et al., 2013b).

H3: Higher levels of self-assessed positive affect will predict higher levels of a) identity fusion and b) perceptions of relational bonds with other participants (Chapter 2).

H4: Stronger physiological response to the artificial ritual will promote greater reflexivity and this effect will increase after a delay (Richert et al., 2005).

3.2 Methods

The design of this study was deliberately intended to be as streamlined as possible including only the most important variables in order to increase the ease of replication and provide a core structure that could be further amended with additional measures being inserted after establishing that the core affective manipulation was effective and an initial test was provided of the hypothesised connections between dysphoric arousal and identity fusion, relational bonds, and prosocial behaviour directed at the ingroup.

Ethics

The initial methodology for this study was proposed by Harvey Whitehouse and Jon Lanman, and was then further developed by the current thesis author in collaboration with Ryan McKay and Jonathan Jong. The current thesis author contributed significantly to the study design, produced all experimental material, and together with a team of RAs conducted all of the experimental sessions.

The Central University Research Ethics Committee (CUREC) reviewed this research using human participants and approval was received for the experimental protocol (SAME/CUREC1A/12-14). All procedures conformed to the ethical codes prescribed by Oxford University. A description of the study and what it involved, together with a consent form was provided to all participants prior to obtaining final agreement to participate. All participants were aware that they could end their participation, without any penalty, at any time, and were given the opportunity before, during, and after the experiment to ask

questions. As the study involved deception, concerning the origin of the ritual sequence and the purpose of the performance, participants were provided with a detailed debriefing sheet upon the conclusion of the experiment and details to contact the experimenters if they had any further questions. The experimental manipulation was intended to generate dysphoric arousal and was anticipated to be unpleasant for participants; however, the negative effects were also designed to be temporary and non-enduring. Appropriate information about counselling services was prepared in the case that any participants reported a severe negative reaction to the experience, but no such strong reaction was reported. All sessions were conducted by the current thesis author, a research assistant, or by the laboratory manager, Dr. Jonathan Jong, at the Pauling Centre of Oxford University.

Recruitment & Sample

135 participants of both genders (55% Female, 45% Male; $M_{\text{Age}} = 23.01$, $SD = 4.298$), were recruited from the general Oxford public using advertisements placed on a free bulletin board (www.dailyinfo.co.uk) and the department of experimental psychology's online list of on-going experimental studies. The study was restricted to ages 18-35, to limit the diversity of demographic factors across groups, and to those without a history of mental illness, due to the dysphoric manipulation. Participants completed the study in groups of 3, with 45 groups completing the first session. For the second session, there were 4 groups in which the full group did not return and the second session had to be abandoned, resulting in 41 groups (123 Participants) completing both sessions. The sample recruited was predominately White (69.9%, $N = 93$), with smaller proportions of Chinese (7.5%, $N = 10$), Mixed (6.8%, $N = 9$), South Asian (5.3%, $N = 7$), Black (5.3%, $N = 7$) and Other (5.3%, $N = 7$) ethnic

backgrounds present. 57.9% ($N=76$) of the sample identified as British (English, Scottish, Welsh & N. Irish) and the remainder were widely dispersed but predominately from within Europe (29.3%, $N=39$). Christians accounted for 31.6% ($N=42$) of the sample and 56.4% ($N=75$) identified as none, there were also small numbers of Hindus (3%, $N=4$), Jews (3%, $N=4$), and Muslims (3%, $N=4$).

Pre-testing of Stimuli

Initially, three audio tracks were created for use in three independent conditions: a dysphoric (baby crying), euphoric (baby laughing), and neutral (baby gurgling) track. The audio tracks were constructed using public domain sound clips downloaded from online sound effect repositories (e.g. www.soundjay.com, www.soundbible.com) and were selected based on the length of the recording (close to 2 minutes), the quality, and the absence of background noise. All clips were edited using Audacity software (version 1.3.4-Beta, <http://audacity.sourceforge.net/>) using the normalize function (set at -3dB) to ensure a consistent maximum level of volume, while retaining the characteristic variations in amplitude of each clip. During the experiment the clips were played from a laptop using surround sound speakers, which were positioned in the corners of the laboratory space and maintained at a consistent volume for all sessions. Unfortunately, no precise measure of volume was collected but the speakers present were set to approximately 75% of the maximum level.

Prior to launch, the three audio tracks were piloted on independent lab members ($N=22$), who listened to one randomly selected track and rated their affective response on semantic affect manikins (Bradley & Lang, 1994), with 9-point rating scales for affective

valence and arousal (higher scores equal more positive and more arousing responses). A one-way ANOVA with affective valence as the outcome measure and the three tracks as the categorical comparison found that there was a significant difference in valence ratings, $F(2,18)=34.862$, $p<.001$. Post-hoc Tukey tests revealed that the mean rating of both baby laughing ($M=7.429$, $SD=1.33$) and baby gurgling ($M=6.71$, $SD=1.60$) were significantly higher ($p<.001$), and hence more positive, than the crying baby stimulus ($M=2.571$, $SD=.534$). However, there was no significant difference in valence between the gurgling and laughing clips ($p=.504$). Moreover, in a second one-way ANOVA this time with arousal as the outcome measure, although a main effect, $F(2,18)=3.267$, $p=.062$, close to significance was found, post-hoc Tukey tests suggested that the difference observed was being driven by the difference ($p=.050$) between the baby crying ($M=6.857$, $SD=1.069$) and baby gurgling stimulus ($M=4.857$, $SD=1.11$). No significant difference was found between the baby laughing ($M=5.857$, $SD=1.621$) stimuli and the other two stimuli (both $p=.425$). Consequently, for Study 1 the euphoric condition was dropped in favour of a two-condition (control vs. dysphoric) design. It was intended however that this condition would be reintroduced along with additional measures in follow up studies if a main difference was found between the dysphoric (crying) and neutral (babbling) condition.

Experimental Procedure

Each experimental group was conducted with three participants and one primary experimenter, with an additional assistant present during training sessions. When fewer than three participants arrived for the session, those who had shown up were given the option to reschedule for another session or to cancel their participation and receive partial

compensation. All participants were informed that they would receive a minimum of £15 for their participation in the two sessions but that they might receive more, depending on certain decisions they made during the study.

Prior to attendance, participants were randomly assigned to either the crying baby (dysphoric) or gurgling baby (neutral) condition. The only difference between the two conditions was in descriptive references to the baby sound effect in the questionnaire and the audio stimulus played during the experiment. The primary experimenter was not blind to the condition, as they were required to select the appropriate audio file and were present throughout the ritual task performance.

The experimental protocol involved following a six-stage process to ensure consistency between groups. A detailed description of each step and the measures included follows below.

Session 1

Step 1. Arrival

In a reception area, participants were welcomed individually by the experimenter. While waiting for the full group to arrive, the experimenter attempted to prevent the discussion of personal details by asking about travel arrangements. When all three participants had arrived, they were welcomed collectively, and provided with study information and consent forms. Participants were given time to read the material and asked to indicate whether they still wished to participate. Recruitment material specified that the experiment required participants not have any prior relationship and that it was therefore

unsuitable to sign up with friends for the same session. The experimenter restated this requirement before each session, however no pre-arranged sign-ups were divulged⁴⁴.

After consent forms were collected, all participants were provided with anonymous IDs and told to use these to identify their questionnaire in the experimental room. All individuals were also told that they were assigned to *Team Ten*, which was intended to provide a minimal group identity that was consistent across all groups but also implied the existence of other teams. Finally, the experimenter provided a brief summary of what the study would involve and how long each section would take.

Step 2. HR Fitting & Baseline collection

After providing the summary, the experimenter then introduced the heart rate (HR) monitoring equipment, demonstrated how to correctly fit the chest transmitters, and asked participants to use the bathroom facilities to fit the bands in private. After all the transmitters were fitted, a matching HR wrist-receiver was provided and participants were lead to the experimental room. Providing the HR equipment at this early stage was intended to allow for some degree of habituation before the actual ritual performance and to enable technical problems to be resolved. In the experimental room, there were three seats all set facing away from each other. Participants were asked to locate the questionnaire with the ID provided to them and to sit at the relevant spot. Once seated, all participants received instructions on how to activate the HR monitors and then, after necessary adjustments were made, baseline measures were recorded.

⁴⁴ Two participants did report familiarity due to attending the same course but they had not intentionally arranged to participate together, and in this case the session was allowed to proceed, as participants were also asked to indicate any prior relationship in the questionnaire.

Step 3. Cover Story & Ritual Sequence

Following the baseline HR measure, participants were asked to turn to the next page of their questionnaire and read the introduction and the details of the ritual sequence, as the experimenter set up the necessary equipment.

Table 3-1: Instructions for Ritual Sequence

Step	Action
1	Pour the powder into the water in the large earthenware bowl.
2	Stir this mixture with your right hand, 25 times, going clockwise.
3	Grind the seeds in the mortar, using the pestle.
4	Pour the ground seeds into the mixture.
5	Put the plant material into the mortar.
6	Stir the mixture with your right hand, 25 times, going anti-clockwise.
7	Immerse your right hand into the mixture, and sprinkle droplets of water from your right hand over the plant material. You may then dry your hands on the towel provided.
8	Pour the small bowl of water into the perforated bowl, and wait until the water has stopped filtering through.
9	Crush the plant material with the pestle.
10	Put the plant material into the mixture.

The 10-step ritual sequence was designed to provide a collective task that would be identical for each individual, included repetitive and causally opaque actions, and would last for about 5 minutes. The ritual sequence involved tactile aspects (e.g. immersing hands in the mixture, sprinkling droplets on the plant, and grinding seeds) to increase participant involvement. The elements were also intended to be reminiscent of food preparation but with arbitrary specifications (e.g. anti-clockwise stirring) to create some ambiguity about

the purpose of the ritual. 10 steps were found during early piloting to be an appropriate length to make the task last the desired time and to remain possible, but challenging, to recall. A variety of natural materials (seeds, red powder, plant material) were included to provide a range of targets for analogical reasoning and reflection (Richert et al., 2005). The full ritual sequence is provided above.

The 10-step ritual sequence was preceded by the cover story for the study. This was identical across conditions except for the title of the ritual ('wailing infant' in the dysphoric condition, 'gurgling infant' in the neutral condition) and read as follows:

"We're studying the rituals of an ancient civilization. In particular, we are interested in a specific ritual that researchers found described on some preserved animal skins [bearing the title 'wailing infant'/'gurgling infant']. Although we have been able to reconstruct a detailed set of instructions about how to perform the ritual, we have very little information about what the ritual was for or what the various actions symbolised. One way anthropologists try to figure out the purpose of particular rituals is to examine people's intuitions about them, having performed them. We will ask you to perform the ritual and then ask you various questions about what it may have meant. [To help increase realism (within the confines of the lab space), we'll be adding a few elements (e.g., sound effects) to partially simulate the environment in which this ritual may have been performed]. It is important that you not communicate with anyone from your group or anyone else's group during the period between sessions, as this will nullify your answers."

Once the materials were prepared on a space on the floor, the experimenter confirmed that all participants had read the material through at least once, and then proceeded to re-read the cover story aloud. The experimenter then indicated where each participant should

sit and in what order they should perform the ritual sequence, explained as being determined by their randomly allocated IDs. A demonstration of each step of the ritual sequence was then provided by the experimenter. Participants were also notified that they could take their instruction sheets with them during their performance and consult them as needed. This was not stated earlier in order to encourage greater attention on the material. Participants were reminded that they should not communicate during the ritual and that, when not performing the sequence, they should watch the other performances carefully. The experimenter explained that they would be present and taking notes during the ritual performance but would not intervene or provide any prompts, including when to change positions. Participants were also informed they could raise their hand to ask for assistance but that they should try and treat the ritual with respect and seriousness. These instructions were intended to reinforce the cover story, encourage engagement with the task, and reduce the influence of the experimenter's presence.

To perform the ritual, participants were asked to sit on the floor, in a triangle formation. One participant then would perform the entire ritual sequence while the other two participants observed. After one sequence was completed the participants switched positions by rotating clockwise and the next sequence was performed. This meant that the ritual was performed a total of three times during each session, once by each participant. The signal to begin the sequence was when the experimenter started playing the background sound. The sound clip was played on a loop until the end of the experiment. Additional details of the properties of the audio tracks are provided in the pre-testing section. Between performances, the experimenter refilled the necessary materials (water,

seeds, red powder, and plant material) and exited the room briefly to empty the contents and clean the main bowl. This took approximately one minute to complete. Participants were requested to use this time to reposition themselves and to remain sitting in silence until the experimenter returned.

Step 4. Post-Ritual Questionnaire

After completing the ritual performance all participants were asked to remain seated and to record three HR measures from their wrist monitor: i) their current heart rate, ii) average heart rate during the ritual, iii) and maximum heart rate during the ritual. Participants took their own measures at this stage, as during piloting it was found to be more disrupting and too time consuming for the experimenter to individually collect the results. Participants then returned to their seats and completed a short questionnaire which included the following measures:

1. The Positive and Negative Affect Schedule, using a 20 item adjective list (PANAS: Watson, Clark, & Tellegen, 1988) that was rated on a five point scale.
2. A brief open-ended question eliciting their thoughts about the meaning of the ritual, intended to capture an element of SER (Richert et al., 2005; Whitehouse, 2004): *We are interested in your intuitions about the meaning of the ritual you performed in the previous session. In the space provided below, please write down your thoughts about the meaning of this ritual, what it may have been intended to do and symbolize.*

3. Four novel measures designed to address relational bonds, specifically the sense of connection and shared meaning, felt between participants employing a 6-point response scale (Not at all - Very Much)⁴⁵:
 - i) *Do you feel like you know any of the other participants well?*
 - ii) *Do you feel a sense of connection with any of the other participants?*
 - iii) *Do you feel an emotional attachment to any of the other participants?*
 - iv) *Do you feel like you shared something meaningful with any of the other participants?*
4. Empathic and Distress response to infant sounds using an eight adjective list, measured on a 5 point scale (Batson, O'Quin, Fultz, Vanderplas, & Isen, 1983) to provide a manipulation check and some control over idiosyncratic responses.
5. Demographic information – age, gender, race/ethnicity, religion, nationality.

Additional measures, including group identification scales, were initially included in this section but removed during preparation in order to reduce the duration of the experiment and keep the initial design streamlined.

Once all participants had finished the questionnaire they were reminded of the date and time for their second session was, asked not to discuss the experiment with other participants or friends, and dismissed.

⁴⁵ The wording of these questions was constructed through consultation with Harvey Whitehouse, Ryan McKay, and Bill Swann with the intention being to address the specific kind of relational bonds predicted in Modes theory. Existing relational measures were found to be unsuitable for the context or to focus on aspects that differed from Modes theory.

Session 2

Step 5. Questionnaire

Following a 3-day delay, participants returned for the second session. Full sessions were only run when all three participants returned, otherwise a shortened session was conducted which involved completing the questionnaire and skipping the resource allocation game. After returning to the same experimental room and seating arrangement, participants individually completed the second questionnaire, which consisted of a repetition of the measures for SER, and relational bonds from session 1 and the following additional items:

- The (7 image) pictorial fusion measure⁴⁶ (Gómez, Brooks, et al., 2011; Swann et al., 2009) with the group circle designated as *Team Ten* (the group identity introduced in S 1).
- Three questions concerning previous relationships with other participants: The other group members were labelled with anonymous designations (e.g. Group member 1) and participants were asked if they had previously met any of the members before participating. For those who answered yes, they were asked to elaborate on i) how long they had known the person before the experiment, ii) how they would describe their relationship (acquaintance, friend, classmate, etc.), and iii) how much they liked the person.

⁴⁶ The pictorial fusion measure was included rather than the verbal scale due to time constraints and the individual items being incongruent with the minimal group setting used in the experiment.

Step 6. Resource Allocation Game (Cooperation) & Dismissal

After all questionnaire were collected, the experimenter provided participants with the instructions for a *resource management and decision making task*. This task was a variant of a public-goods game (Croson & Marks, 2000; Hirshleifer, 1983), recently used by Wiltermuth & Heath (2009: Study 3), and was played over five rounds. In the game, each participant was allocated 10 tokens during one round and had to decide how much to contribute to a public (*Team Ten*) account and how much to keep in their private account. At the end of each round, each participant would receive 10p for each token placed in the *Team Ten* account and 20p for every token in their private account. Tokens in the public account earned 10p for all participants, whereas tokens in private accounts only added value to the individual.

In this type of game, individuals have an incentive to free-ride, as they can obtain more direct value from keeping their tokens in their private account, and benefiting from others' contributions to the public account. Alternatively, if all participants were to donate their tokens to the group account then this maximizes the group's potential earnings. The game thus intentionally reflects a 'tragedy of the commons' scenario (Hardin, 1968), where selfishness is more beneficial to any individual, but cooperation would benefit the larger group, more equitably. Due to concerns about floor effects, stemming from the minimal nature of the experimental group involved, feedback was provided regarding group contributions to the *Team Ten* account, following each round. But participants did not learn the contributions of individual group members. This collective feedback methodology has been found to lead to higher overall levels of cooperation (Cason & Khan, 1999; Sell &

Wilson, 1991; Weimann, 1994). To provide motivation participants were also informed that, at the end of the experiment, they would receive a real bonus payment based on the allocations made from a randomly selected round and should thus treat each round as the potential pay-out.

Participants read the instruction sheet individually, listened to the experimenter summarising the task, and then completed two practice scenarios to ensure that they had understood the rules. The task only proceeded when all group members had answered the practice questions correctly and confirmed they understood the task. During the rounds, participants were not permitted to speak or communicate with each other and were asked to raise their hand if they had a question for the experimenter. The experimenter verbally introduced each round and, after a delay, collected the answer sheets. The combined contributions to the *Team Ten* account for the round were then calculated, written on each participant's answer sheet, and then returned before the start of the next allocation round.

After five rounds were complete, a random payment round was selected for each individual, and they were provided with their payments (£15 for 2 sessions + earnings from one round of the PGG). A collective debriefing was then provided about the purpose of the study and the artificial nature of the ritual they performed. The experimenter answered any additional questions and then dismissed participants, reminding them not to discuss the nature of the experiment until the study period was finished.

Materials

Heart rate measures (avg. and max) were collected using 3 x coded Polar FT2 heart rate monitor watches, with matched T31 chest band transmitters.

The physical materials for the ritual sequence comprised of:

- 1 x large unfired earthenware bowl⁴⁷, filled 1/3rd with water.
- 1 x mortar and pestle
- 4 x small glass containers that held i) plant material (leaves collected from a nearby park), ii) an odourless red dye, iii) chia seeds and iv) water
- A water timer made from: 1 x small bowl with a perforated base (for slowly releasing water), 1 x medium curved vertical pot (for collecting the water)
- 1 x large jug (for refilling water)
- 1 x wooden board

3.3 Results

In all results reported below, unless otherwise stated, 9 participants were excluded due to having previous relationships with other participants, being outside the stipulated age range, guessing the goal of the study, or failing to complete substantial portions of the questionnaire/protocol. Responses to the open-ended SER question on the meaning of the ritual, indicated that 21 participants were suspicious about the cover story or guessed something close to the general purpose of the study, however all findings reported below were unchanged when these responses were excluded and hence they were retained in the analyses below.

⁴⁷ All earthenware products were sourced from the *Objects of Use* store in Oxford and were selected based on the simple clay materials used. The intention in using such materials was to make the objects appear distinctive and avoid identifying features to keep their origin ambiguous.

Data Preparation

Prior to analysis, the demographic composition of the two conditions was examined for any significant differences in the age, gender, or religious background but no significant differences were detected, using student t-tests.

Relational bonds

A principal components analysis (PCA) was conducted on the responses from session 1 to the four novel items intended to measure relational bonds with other participants (*feeling of knowing, sharing a connection, sharing something meaningful and sense of emotional attachment*), each of which was measured on a 6-point scale. The suitability of a PCA was assessed prior to analysis. Inspection of the correlation matrix showed that all variables had at least one correlation coefficient greater than .3. The overall Kaiser-Meyer-Olkin (KMO) measure was .804, with individual KMO measures all greater than .7, classifications of 'middling' to 'meritorious' according to Kaiser & Rice (1974). Furthermore, Bartlett's Test of Sphericity was statistically significant ($p < .001$), indicating that the data was likely factorizable.

The resulting PCA identified a single component, which accounted for 67.66% of the total variance. Visual inspection of the scree plot confirmed that a single component should be retained (Cattell, 1966). In addition, a single-factor solution was a coherent outcome as the measures, although addressing sense of shared meaning and attachment, had been intended to examine the concept of *relational bonds*. The four items also displayed a high reliability when combined into a scale (S1, $\alpha=.832$ and S2, $\alpha=.881$).

Positive / Negative Affect

Responses to the 20-item PANAS measure was analysed according to established protocols (Watson et al. 1988), which involved aggregating the responses from the 10-positive items ($\alpha=.887$) and the 10-negative items ($\alpha=.851$) into independent positive and negative affect scores.

Cooperation / Prosociality

For the Public Good's Game, contributions to the group account across all five rounds were summed and then subtracted from the sum of contributions to the participant's private account across all five rounds. This was used as a measure of cooperation, in line with the existing experimental literature (Croson & Marks, 2000; Hardin, 1968; Hirshleifer, 1983).

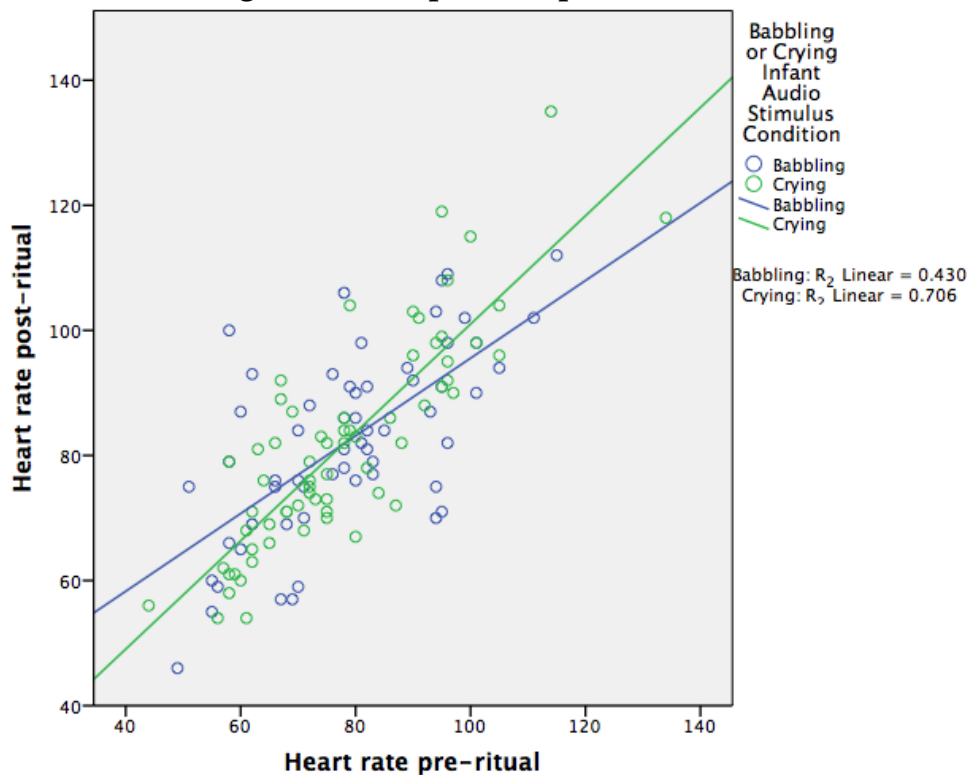
Manipulation Checks

To examine whether the experimental manipulations were successful in producing distinctive affective responses, both physiological and self-report measures were assessed. For physiological response, the three post-manipulation HR measures were examined in a series of independent ANCOVAs⁴⁸. In these models, the post-manipulation HR measures were independently entered as outcomes, experimental condition was included as a predictor, and baseline HR was controlled for as a covariate. In all the HR analyses reported, a total of 28 responses were excluded due to errors with the HR monitoring equipment. Controlling for the baseline HR measure, experimental condition was not

⁴⁸ There was a linear relationship between pre- and post-HR measures for each experimental condition for all three measures, as assessed by visual inspection of a scatterplot. Homogeneity of regression slopes was also found, as assessed by preliminary linear regressions that included interaction terms, except in the case discussed.

found to be significantly predictive of any difference during the ritual performance in either *average* heart rate, $F(1,104)=1.632$, $p=.204$, or *maximum* heart rate, $F(1,104)=.003$, $p=.954$. However, for the average HR post-ritual measure, when testing for the assumption of homogeneity of regression slopes, a main effect approaching significance for experimental condition was found, $F(1,102)= 3.593$, $p=.061$, along with a significant interaction between condition and baseline HR, $F(1,102)=4.190$, $p=.043$. Although this violated the necessary assumptions for an ANCOVA and makes the main effect observed potentially unreliable, it also suggests that there was a distinctive physiological response between conditions with HRs increasing at a greater rate in the crying baby condition (Fig.3.1).

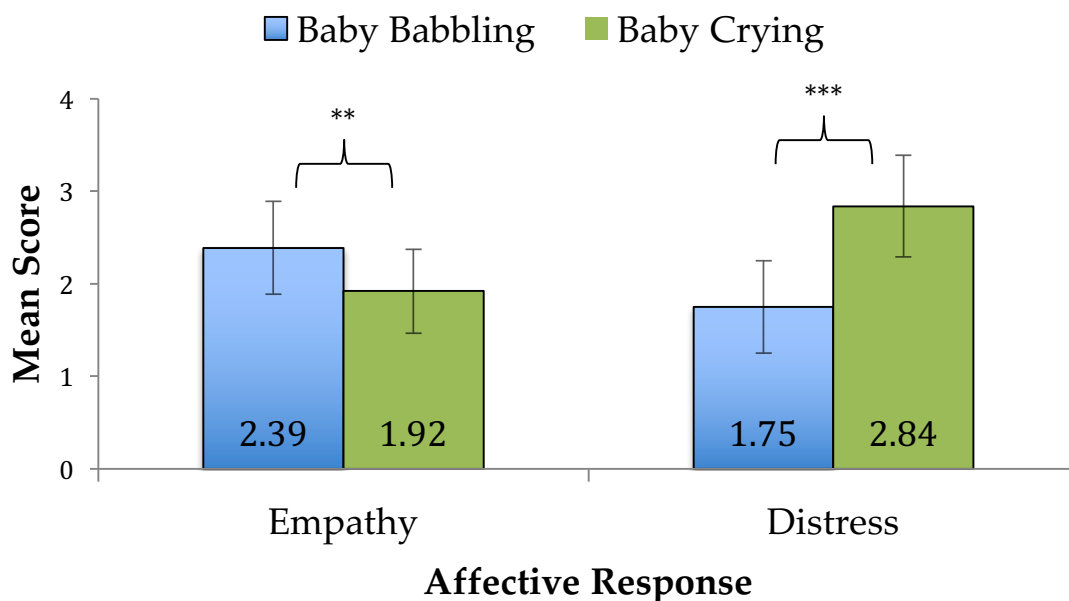
Figure 3-1: HR pre- and post-ritual



In summary, the physiological heart rate evidence is ambivalent; two of three measures showed no significant difference and the single measure in which a significant difference was found was the measure taken *after* the ritual performance.

A clearer difference between conditions was found when comparing Batson et al.'s (1983) self-report measures of empathy ($\alpha=.902$) and distress ($\alpha=.894$), in response to the baby audio stimulus (Fig.3.2). The babbling baby elicited a more empathetic response ($M=2.391$, $SD=1.063$) than the crying baby stimulus ($M=1.915$, $SD=.972$); $t(124)=2.617$, $p=.010$, and conversely, stronger distress was experienced in response to the crying baby ($M=2.835$, $SD=1.150$) over the babbling baby stimulus ($M=1.750$, $SD=0.898$); $t(124)=5.914$, $p<.001$.

Figure 3-2: Affective response to baby sounds stimulus



Note: Error bars represent standard deviations.

Affective response between conditions were also compared using the aggregated positive ($\alpha=.887$) and negative ($\alpha=.851$) self-report responses on the PANAS (Watson et al., 1988). On these measures, there was no significant difference found on positive affect,

$t(120)=.867, p=.388$, but there was a difference approaching significance on negative affect, $t(123)=1.888, p=.061, d=0.340$, with those in the crying baby condition reporting more negative affect ($M=16.115, SD=7.134$), than those in the babbling baby condition ($M=14.094, SD=4.628$).

The self-report measures thus provided stronger support for the success of the manipulations. The crying baby condition generated more distress in participants and this seems to have produced a slight increase in negative affect. The lack of any detectable difference in positive affect is potentially due to all participants experiencing relief that the challenging ritual performance was complete.

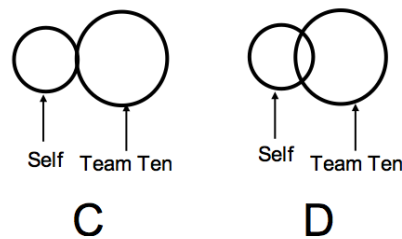
Hypothesis 1 (Crying condition = Higher Fusion, More Relational Bonds)

A two tailed t-test of ratings on the 7-point pictorial fusion measure was conducted to assess whether the dysphoric ritual performance was associated with higher levels of identity fusion with the group. A t-test analysis was chosen despite a K-S test indicating a non-parametric distribution, $D(145)=.145, p<.001$, as the skew and kurtosis of the fusion responses were limited and visual inspection of Q-Q plots and histograms suggested an approximate parametric distribution.

The results obtained indicated that there was no significant difference between the levels of identity fusion in the crying ($M=3.64, SD=1.669$) and neutral ($M=3.45, SD=1.688$) conditions, as assessed by the 7-point visual fusion measure, $t(115)=.631, p=.529$. Furthermore, in both conditions, the most commonly endorsed images (46.6% of responses) were option C, which portrays no overlap, and option D, which represents a minor overlap between the self and other icons. Only 10 participants endorsed option G, which is used to

represent a fully fused state. Given the novel nature of the groups, this is not a surprising result.

Figure 3-3: Most common fusion images endorsed in both conditions



Next, a non-parametric Mann-Whitney U test was used to compare conditions for the combined measure of relational bonds at both time periods. Again, no significant differences in relational bonds between conditions was detected in both sessions: S1- $U=1,986$, $z=.010$, $p=.992$; S2- $U=1,847$, $z=.583$, $p=.560$. Furthermore, there was also no difference in the change in relational bonds between S1 and S2: $U=1,723$, $z=-.090$, $p=.928$. Non-parametric comparisons were used due to the relational tie measures being moderately positively skewed, appearing far from normal in Q-Q plots, and obtaining results indicating non-parametric distributions from K-S tests for both sessions: S1- $D(118)=.134$, $p<.001$, S2- $D(118)=.164$, $p<.001$.

Table 3-2: Descriptive statistics for identity fusion and relational bonds

Variable	Skew (SE)	Kurt. (SE)	α	M (SD)
Pictorial Identity Fusion	-.444 (.224)	-.328 (.444)	-	3.550 (1.674)
Relational bonds (Session 1)	-.716 (.223)	-.181 (.442)	.832	1.040 (.793)
Relational bonds (Session 2)	-.592 (.102)	.420 (.204)	.881	.873 (.779)

Although no main difference between conditions was found the theorised connections between identity fusion and relational bonds was examined through an additional correlation analysis on the identity fusion and relational bonds measures. Using

spearman's rho to control for non-parametric distributions, significant positive correlations were observed between the pictorial fusion measure and relational bonds measures in S1; $r_s=.283, p=.002$ and this was found to increase S2; $r_s=.512, p<.001$.

Hypothesis 2 (Crying Condition = More Cooperation)

Prior to analysis of the PGG task, in addition to the 9 exclusions due to pre-existing relationships, 7 responses were excluded because the participants either communicated with each other prior to the task, discussing the ideal strategy to adopt, or reported not understanding the task during debriefing. Furthermore, four groups did not complete the allocation task in the second session due to some group members not returning. Consequently, the total number of responses analysed for H2 was $N=111$.

Table 3-3: Descriptive statistics for cooperation

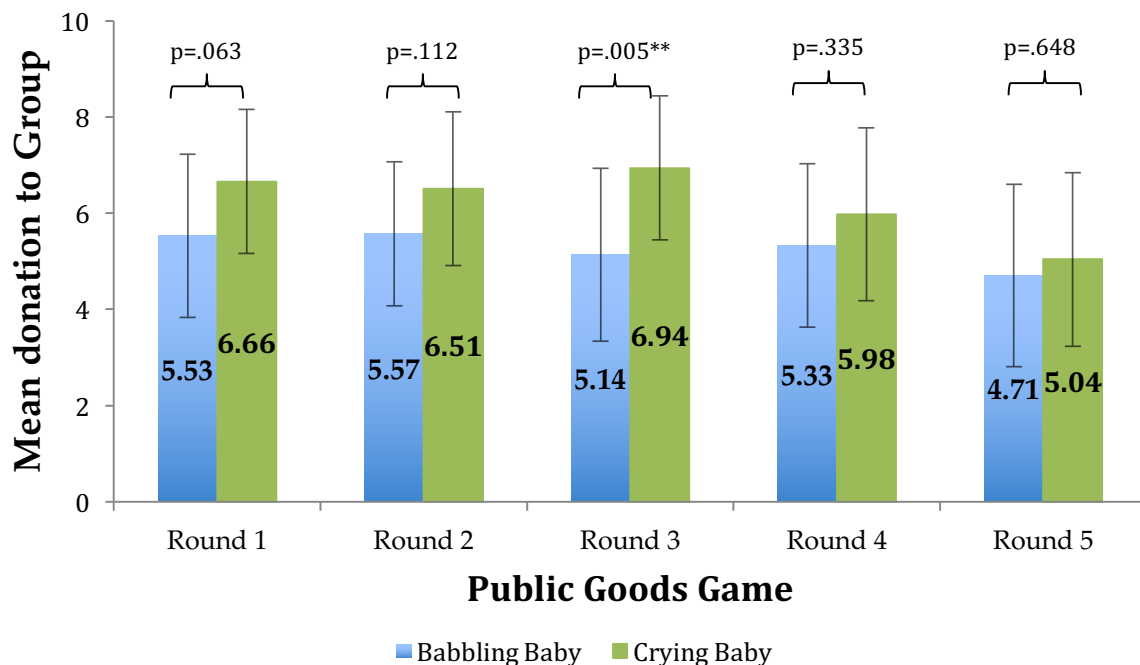
Variable	Skew (SE)	Kurt. (SE)	α	M (SD)
Cooperation in PGG	-.047 (.229)	-.480 (.455)	-	7.189 (26.766)

Examining these responses, a K-S test indicated a non-normal distribution for the cooperation measure; $D(111)=.197, p=.003$, but levels of skew and kurtosis were acceptably low (Table 3.4) and a visual inspection of Q-Q plots and histograms, suggested only minor deviations from normality. Consequently, parametric tests were employed.

When examined across all rounds, the difference in cooperative donations between conditions indicated a marginally significant trend, $t(109)=1.933, p=.056, d=0.369, M_{diff}=9.712$, 95% CI [-19.671, .245]. The direction of the difference was also in accordance with predictions; with average cooperation in the crying baby condition ($M=12.264, SD=23.567$), higher than in the babbling baby condition ($M=2.551, SD=28.812$). For each of the five

rounds, the average group contribution was higher in the crying baby condition (Fig.3.4), but the difference only reached significance on round 3.

Figure 3-4: Cooperation in the PGG between conditions



Note: Error bars represent standard deviations.

These results are suggestive of higher overall cooperation in the dysphoric condition but are also far from conclusive. To further explore whether any individual factors potentially promoting greater levels of cooperation could be identified exploratory spearman's correlation analyses were conducted between the cooperation measure and demographic factors (age and sex), positive and negative affect (PANAS), empathy and distress, relational bonds, and fusion measures (Table 3.4). From these measures only positive affect and relational bonds were found to have a significant positive correlation with the cooperation measure.

Importantly, the correlation observed between cooperation and relational bonds remained significant in both sessions when controlling for positive affect scores, with

bootstrapping (1,000 samples) applied to help address non parametric distribution: Relational bonds S1, $r=.226$, $p=.017$, BCa CI [.031, .395]; Relational bonds S2, $r=.188$, $p=.048$, BCa CI [.002, .359]. Although the p-values were statistically significant in both sessions and this increases confidence in the validity of the finding, the correlation in S2 was just below the commonly applied cut off point and thus some caution is warranted to avoid over interpreting this finding. Alternatively, the reduced correlation could also be that the relational bonds reported during S2 were less relevant to the shared ritual experience.

Regarding the observed correlation between positive affect in S1, it was only recorded in this session, and the cooperation measure the correlation was found to be non-significant, $r=.143$, $p=.135$, BCa CI [-.016, .299] when the relational bonds measure from S1 was controlled for. This suggested that the relationship between relational bonds and enhanced cooperation was more robust than the connection with positive affect. Positive affect was also weakly correlated with relational bonds across both sessions, S1- $r_s=.349$, $p<.001$; S2- $r_s=.320$, $p<.001$.

Table 3-4: Spearman's correlations with cooperation in PGG

Variable	Visual Fusion	Relational bonds (S1)	Relational bonds (S2)	Positive (PANAS)	Negative (PANAS)	Empathy	Distress
Cooperation	.153	.215*	.236*	.185*	.042	-.040	.027

$p<.05^*$

These connections are examined further in H3 but both relational bonds and positive affect are poor candidates to account for the difference observed between conditions as there was no significant difference between conditions on either of these measures.

Hypothesis 3a (Positive Affect = Higher Fusion)

Based on the results from the BJJ Study reported in Chapter 2, it was anticipated that subjective positive assessment of the ritual experience, regardless of condition, would be correlated with higher levels of identity fusion. To examine this, simple Pearson correlations were conducted as the variables displayed parametric distributions (Table 3.6) between the visual fusion measure and the positive and negative affect measures. However, no significant correlation was found with fusion for either the positive or negative measure (all $p > .273$).

Table 3-5: Descriptive statistics for PANAS measures

Variable	Skew (SE)	Kurt. (SE)	α	M (SD)
Positive (PANAS)	.483 (.219)	-.045 (.435)	.886	25.213 (8.091)
Negative (PANAS)	-.444 (.224)	-.328 (.444)	.851	15.080 (6.045)

Hypothesis 3b (Positive affect = Relational Bonds)

Subjective positive assessment was also anticipated to be associated with stronger relational bonds between participants. Correlation analyses conducted in H1 (Table 3.3) demonstrated that this was the case with a positive correlation between relational bonds and positive affect observed in the measures collected in both sessions. However, positive correlations were also observed with the relational bonds and the negative affect and empathy/distress measures. To investigate the relative strength of these associations two hierarchical linear regression analyses, involving two stages, were conducted. The first model had the relational bonds measure from session one as the outcome and the second model used the relational bonds measure from the second session. In both models the same hierarchical structure was employed: in the first stage, the empathy and distress measures

were entered along with negative affect score, demographic factors were excluded due to no observable correlation or specific theoretical prediction. In the second stage, subjective positive affective assessment scores were added to produce the full model.

Testing the variables and the two models generated for assumptions: first, examination of standardised residuals identified only a single case as an extreme outlier, designated as a value greater than 3 SDs. Further examination of leverage and influence, using Cook's and Mahalanobis distance values, indicated no other problematic cases. Accordingly, the single outlier was removed for both models and analyses were rerun. Both models demonstrated independence of residuals, confirmed through Durbin-Watson statistic values approaching 2. Partial regression plots also indicated suitably linear relationships for all variables and there was no indication of multicollinearity as assessed through examination of correlations and VIF and tolerance statistics. Finally, examination of the residuals in a histogram, P-P plot, and Normal Q-Q plot indicated an appropriately normal distribution.

Table 3-6: Hierarchical multiple regression on relational bonds S1

Variable	Outcome: Relational bonds S1			
	STAGE 1		STAGE 2	
	B (SE)	β	B (SE)	β
Constant	.362 (.230)	-	-.008 (.283)	-
Empathy	.168 (.066)	.224*	.122 (.068)	.162
Distress	.191 (.072)	.283**	.155 (.073)	.229*
Negative (PANAS)	-.009 (.014)	-.072	-.005 (.014)	-.039
Positive (PANAS)	-	-	.019 (.009)	.203*
R^2	.114		.149	
F	4.966**		5.040***	

*N=119, *= $p<.05$, **= $p<.01$, ***= $p<.001$, shading added to emphasis variables with significant coefficients.*

With relational bonds in session 1 as the outcome, the full model (Table 3.7) was able to predict a limited portion of the variance in relational bonds; $R^2=.149$, $F(4,115)=5.040$, $p<.001$.

Although the addition of positive affect did result in a significant improvement over the initial model, the increase in explanatory power was minor; $\Delta R^2=.035$, $F(1,115)=4.779$, $p=.031$. Yet it is also notable that, counter to observed correlations, there was no significant independent contribution evident for the negative affect measure. This is likely due to the distress measure capturing a substantial portion of negative affect, an interpretation supported by the moderate correlation observed between the two measures ($r=.577$, $p<.001$).

Table 3-7: Hierarchical multiple regression on relational bonds S2

Variable	Outcome: Relational bonds S2			
	STAGE 1		STAGE 2	
	B (SE)	β	B (SE)	β
Constant	.192 (.236)	-	-.198 (.286)	-
Empathy	.139 (.067)	.190*	.090 (.069)	.123
Distress	.157 (.073)	.237*	.123 (.073)	.185
Negative (PANAS)	.001 (.014)	.011	.005 (.014)	.038
Positive (PANAS)	-	-	.021 (.009)	.222*
R^2	.098		.141	
F	3.916**		4.396***	

$N=111$, * = $p<.05$, ** = $p<.01$, *** = $p<.001$, shading added to emphasis variables with significant coefficients.

In the next model with relational bonds in session 2 as the outcome, the full model (Table 3.8) was able to predict a limited portion of the variance in relational bonds; $R^2=.141$, $F(4,107)=4.396$, $p=.002$. Again, the addition of positive affect in stage 2 resulted in a significant but small improvement in explanatory power; $\Delta R^2=.043$, $F(1,107)=5.361$, $p=.022$. In this model however none of the other variables contributed significantly after the inclusion of the positive affect measure. This suggests that subjective assessment of positive affect was the strongest predictor from the variables modelled and did not seem to decay in

predictive ability after a 3-day delay. The similar patterns observed in models for both outcomes also increases confidence in the validity of the findings.

The results from these regressions collectively lend support for the hypothesised connection between subjective positive assessment of ritual experiences and relational bonds. However, the expected association with higher levels of fusion was not observed.

Hypothesis 4 (Physiological Response = More Reflexivity)

To assess the level of reflection that participating in the artificial ritual generated, two hypotheses-blind, independent coders, were asked to assess to rate the reflexivity of the participants open-ended answers to the question about their ‘thoughts about the meaning of the ritual’. Both coders received the same simple instructions and responses were rated on a 7-point scale (1- not reflexive at all to 7- extremely reflexive).⁴⁹ Based on codings of 125 responses, the correlation between ratings for session 1 was moderate, $r=.748$, $p<.001$, with inter-rater agreement (Landers, 2015), calculated using intraclass correlations (Shrout & Fleiss, 1979), indicating a high level of agreement; $ICC(2,2)=.821$, 95% CI [.746, .875]. Ratings for session 2 displayed a weaker correlation, $r=.582$, $p<.001$, and a lower level of inter-rater agreement; $ICC(2,2)=.668$, 95% CI [.528, .766] but the agreement overall was sufficient to enable the scores of both coders to be combined for each session in order to construct two overall *reflexivity scores* for each participant (Table 3.9).

Table 3-8: Descriptive statistics for PANAS measures

Variable	Skew (SE)	Kurt. (SE)	α	M (SD)
Reflexivity (S1)	.574 (.217)	-.537 (.431)	-	2.407 (.129)
Reflexivity (S2)	-.379 (.217)	-.613 (.430)	-	2.104 (.118)

⁴⁹ This single measure of reflexivity was based on wording recommended by Harvey Whitehouse (personal correspondence) and the design of this study was finalised prior to the BJJ survey, hence the difference in measures.

Counter to the findings of Richert et al. (2005), the predicted correlation between stronger physiological response, measured using three HR metrics, and reflexivity was not observed. However, as with the BJJ survey it should be noted that reflexivity was not assessed using the same linguist measures as employed in Richer et al. (2005). Post-hoc analysis found no significant correlations between reflexivity score and self-reported measures of affect in the post ritual period (Table 3.10). There was also no significant difference in reflexivity scores when tested based on conditions, for both sessions; S1- $t(123)=.442, p=.659$; S2- $t(124)=-.223, p=.885$. Additionally, when examined longitudinally using paired t-tests there was no significant change detected in scores between sessions, $t(123)=1.684, p=.095$. This result was also replicated in exploratory subgroup analyses for each individual condition: control- $t(62)=1.338, p=.186$; crying- $t(60)=1.014, p=.315$. In both conditions there was no statistical significance difference detected but the mean reflexivity scores were higher in both condition during the first session: control- S1 $M=2.452, SD=1.555$, S2 $M=2.064, SD=1.418$; crying- S1 $M=2.361, SD=1.304$, S2 $M=2.139, SD=1.249$.

Table 3-9: Spearman's correlations for reflexivity and physio/affect measures

Variable	HR change	Max HR (ritual)	Avg. HR (ritual)	Positive (PANAS)	Negative (PANAS)	Identity Fusion
Reflexivity S1	.063	-.049	.016	.166	.033	.195*
Reflexivity S2	-.091	.068	-.016	.208*	.169	.044

*= $p<.05$

Post hoc analysis also found that higher scores on the visual identity fusion measure were associated with greater reflexivity but this relationship was not observed in the second session (Table 3.10). A more robust relationship was detected with positive subjective assessment which was just outside significance in session 1 ($r=.166, p=.070$), and

displayed a stronger correlation following a delay ($r=.208$, $p=.022$). No such relationship was found with negative affect, and this was further supported by an additional examination of correlations with the distress measure which revealed no significant relationships.

In summary, counter to expectations, stronger physiological response, measured using HR metrics, display no relationship with increased reflexivity on the ritual's meaning. However, in accordance with findings from Chapter 2, greater reflection was associated with more positive affective responses to the ritual performance. Some evidence for the relationship predicted in Modes theory (Whitehouse & Lanman, 2014) between reflection and identity fusion was also found but counter to expectations (Richert et al., 2005) it decayed, rather than strengthened, following a delay.

3.4 Discussion

Overall, the results from this study provided mixed support for existing predictions. Comparing conditions, there was no support found for the prediction that the crying condition resulted in higher levels of identity fusion (H1a), or stronger relational bonds (H1b). However, there was evidence that identity fusion was connected with perceptions of relational bonds. This association was stronger in session 2, but rather than representing a growth over time it is probable that this derives from the fusion measure being taken in the second session. Regardless, the result provides further affirmation of the relational character of the bonds involved with identity fusion.

Despite this, and in accordance with recent field studies (Xygalatas et al., 2013b), greater cooperative behaviour was observed in the dysphoric ritual condition in a five round public goods game (H2). This suggests that dysphoric arousal may serve a role in ritual as a

mechanism for promoting cooperation, similar to synchrony (Wiltermuth & Heath, 2009). This fits with the predictions of both Modes (Whitehouse & Lanman, 2014) and costly signalling theories (Bulbulia & Sosis, 2011) but the trend observed was only marginally significant ($p=0.56$), which reduces confidence in the validity of the finding despite the consistency of the effect across rounds. The significance of such a marginal result is enhanced by considering that participants were complete strangers with no significant affiliation to each other and that the cooperation measure was collected three days after a relatively weak affective manipulation⁵⁰.

Additionally, using PANAS self-report measures, subjective assessments of negative affect were not found to be associated with increased cooperation but positive assessments did display a weak correlation ($r=.236$, $p=.030$). This echoes the finding from the BJJ data presented in chapter 2 and suggests that positive interpretations of dysphoric experiences are important to generating their proposed bonding effect. However, as this positive affect scores were not found to significantly differ between conditions this also seems to be an unsuitable candidate to account for the between condition trend observed. Empathy and distress responses to the baby sounds provided a clearer distinction between conditions but these were also not found to have any significant correlations with cooperation.

Exploring the role of self-assessed positive affect in a ritual context, there was no association found with elevated levels of identity fusion (H3a), but there was evidence that more positively judged experiences were correlated with stronger perceptions of relational bonds with other participants (H3b). A correlation which was also replicated with the

⁵⁰ As assessed by the mixed results from manipulation checks.

repeated measure of relational bonds taken in the second session after a three-day delay, when controlling for positive affect.

The lack of any association with fusion is puzzling, especially in light of the consistent relationship found in the BJJ promotion data. However, there is a clear qualitative difference in the promotion experiences of martial artists and their training schools, and those of a group of strangers performing an unusual task in a lab together. Furthermore, although identity fusion is now recognised as a form of group bonding that is continuous rather than dichotomous, it was originally constructed to account for connections experienced by individuals who are deeply attached and committed to their group (Gómez, Brooks, et al., 2011; Swann et al., 2009). Therefore, it may be unrealistic to expect that even mild forms of fusion could be generated in a novel group setting. However, null results from any individual experiment cannot be used to draw broad conclusions, and therefore further discussion of this issue is reserved for examination of the results in the second study.

The regression analyses that explored the affective predictors of relational bonds, affirmed the results from the BJJ promotion data concerning the importance of subjective positive affect. Although positive affect scores did not differ significantly between conditions, the distinctive responses in empathetic and distress responses to the baby sounds recorded between conditions indicate that the source of the positive response may have been different in each condition. In the control babbling condition there was significantly greater empathy felt by participants towards the infant, whereas in the dysphoric crying condition participants demonstrated an elevated distress response. Thus,

it is likely that the positive affect in the dysphoric condition was prompted by the cessation of the unpleasant stimulus, whereas in the control condition the positive affect could have been due to enjoyment of sharing an unusual experience. According to Batson et al. (1987) distress and empathy trigger different motivational pathways and these can both independently promote affiliative and prosocial behaviour.

The regression analyses on the relational bonds measure from both sessions lend credence to this interpretation as they found empathy and distress to contribute independently. Furthermore, distress remained a significant predictor for relational bonds in session 1 after the positive affect measure was added and in session 2, however after positive affect was added there was only a non-significant trend observed ($p=.097$). This implies that distressing responses promoted relational bonds through some pathway independent from post-event positive affect. In contrast, this does not appear to be the case for empathic responses, as their contributions to the model were superseded in both sessions by the addition of the positive affect measure.

A point of caution here is that the majority of studies exploring the differential impacts of empathy and distress have focused on behaviour and attitudes directed towards the source of the emotional response. However, in this experiment the relational bonds and cooperative behaviour measures were targeting other participants, rather than the crying or gurgling infant. This might help account for the relatively low amount of variance in relational bonds that the regression models accounted for and also for the finding that distress and empathy ratings displayed no association with cooperation (H2).

However, the finding that a shared experience of distress promotes relational bonds, accords well with the account in Modes theory concerning the relational nature of bonds forged in dysphoric rituals (Atkinson & Whitehouse, 2011). Similarly, there is independent research that negative experiences increase desire for social affiliation (Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001; Baumeister & Leary, 1995) and can prompt “tend or befriend” behaviours (Cardoso, Ellenbogen, Serravalle, & Linnen, 2013; Taylor, 2006). Importantly, the lower empathy scores in the dysphoric condition only relate to attitudes towards the infant, participants may have felt increased empathy for their fellow group members but this was not measured directly.

In relation to predictions about reflexivity, counter to Richert et al. (2005), stronger physiological responses (assessed using HR measures) did not result in greater reflexivity about the experience (H4). There was also no increase in reflexivity observed after a delay. The difference in physiological measures of arousal (Galvanic Skin Response was used in the previous study) and the methods used to generate reflexivity measures could account for the discrepancy or, alternatively, it may be that the ritual was not sufficiently arousing or interesting enough to invite significant post-event reflection. This interpretation is partly supported by the overall low mean level of reflection (measured on a 7-point scale, ranging from 0-6) in the full sample; S1 $M=2.407$, $SD=1.432$; S2 $M=2.104$, $SD=1.328$. Additionally, Richert et al. reported significantly greater levels of fright and surprise in the high arousal condition which suggests the ritual experience used was more unsettling for participants. However, it is difficult to assess the significance of this distinction as the self-report measurement employed (summing individual ratings from 1-5 for self-identified points of

fright and surprise) meant that the individual scores would be vulnerable to influence from outliers, a problem further exacerbated by the small sample employed. Furthermore, despite such differences it is important to note that there was no difference detected between experimental conditions on reflection. The lack of personal consequentiality in the present study could also have been a factor reducing motivations for reflection but this should have also applied for the Richert et al. study, which similarly employed participants with no previous relationship and an artificial ritual performance.

21 participants used their open response to provide speculations about the purpose of the experiment and the use of heart rate monitors, rather than reflecting on the elements of the ritual sequence. This tendency was noted in early sessions by experimenters and verbal instructions were changed to emphasise that the question was specifically asking referring to the ritual performance, rather than the purpose of the study. Subsequently, later sessions had much less of this content but it is likely that this reduced the accuracy of the measure. It also suggests that at least some participants were not strongly engaged by the ritual task. This was evident in some sessions during the portions of the ritual sequences that the participants observed rather than performed. Although emphasis was placed on the need to concentrate on the others' ritual performances, a number of participants appeared distracted during the ritual sequence repetitions. This tendency was further exacerbated by the unpleasant audio stimulus in the dysphoric condition, which made participants generally more agitated and prone to fidgeting.

Methodologically, the study affirmed the potential for an artificial ritual paradigm, with only a limited number of participants expressing scepticism about the authenticity of the

ritual and ritual performances typically being completed with care. However, it was also evident that there were significant limitations with the current design and that various components needed to be revised and improved. The main issues are identified below along with brief recommendations about how they could be resolved in study 2.

First, the affective manipulation proved to be weaker than anticipated, although significant differences were detectable between conditions. The audio stimulus was impossible to ignore, although some degree of habituation over time was observed, but it did not fully succeed in generating an integrative atmosphere with the ritual sequence. Additional efforts are needed to increase the strength of the affective manipulation and to better integrate it with the ritual performance.

Second, the ritual task proved to be less *collective* and *interactive* than intended. In particular, the lack of any interaction between participants during each ritual sequence meant that rather than simulating a collective ritual, the performances felt more like an individual ritual repeated three times. This sensation was further exacerbated by the experimenter emptying all of the material prepared between sequences and returning all of the materials and equipment to the original starting state. Additionally, the presence of ritual instruction sheets proved to be more distracting than anticipated and a number of participants looked directly at the instructions for the duration of their performance. This not only reduced opportunities for making eye contact with other participants but also made the performances less engaging for those observing. Finally, the interruption of the experimenter to take the ritual equipment for cleaning made their presence salient and interrupted the flow of performances. The second study should thus seek to reduce

experimenter involvement, increase the collective and interactive elements of the ritual performance, and provide a less distancing method for reminding participants of the ritual steps.

Third, some measures were incongruent with the cover story provided, specifically the public goods game. The cover story provided a convincing explanation for most of the items included in the questionnaire, but several participants expressed suspicion during debriefings concerning the purpose of the economic game and remarked that it did not fit the general tone of the experiment. A smaller number of participants also spontaneously commented that the task appeared to be designed to measure generosity. The current design thus risks undermining the cover story and being too transparent about the purpose of the economic game, potentially leading to responses influenced by demand characteristics. For the second study greater care should be taken to integrate the behavioural measure with the cover story and a less transparent measure of prosocial tendencies should be employed.

Fourth, although each session followed a scripted procedure, there was too much scope for individual variation in sessions, specifically in regards to demonstrations by the experimenter of the ritual sequence. During the demonstration, participants were seated in different positions which meant they had different viewing angles and, furthermore, certain groups interrupted demonstrations with questions and/or asked for certain sections to be repeated. After the first few sessions, experimenters provided an additional warning that the demonstration would only be provided once but this did not always prevent requests for clarification, or interruptions. Furthermore, each experimenter had subtle

differences in their mannerisms and speech, which made it impossible to entirely standardise demonstrations. Although this may not have had a major influence over results it introduces an element of variation that should be avoided if possible. As such, in study 2 instructions for the ritual sequence should be provided through a standardised instructional video, rather than a personal demonstration.

Fifth, all outcome measures collected in the study rely on explicit self-report and behavioural choices. Although this is commonplace in social science studies, given the artificiality of the ritual performance and the low effect sizes observed it would be prudent to include implicit measures (Greenwald, McGhee, & Schwartz, 1998; Sriram & Greenwald, 2009) that can address so-called hot cognitive processes (Abelson, 1963; Brand, 1985; Evans, 2008; Kahneman, 2011). This can help to reduce concerns over demand characteristics and potentially detect changes in attitudes that operate below the level of conscious decisions. Additionally, given the marginal nature of the effects detected in the key outcome variables, study 2 should also aim to collect these measures immediately after the affective manipulation to better detect potentially transient effects.

Sixth, the means used to establish a minimal group identity, namely informing participants that they were part of *Team Ten* and mentioning the existence of other teams, seemed to draw little attention and a few participants expressed confusion over the ‘team’ aspect during debriefings. Earlier drafts of the experimental protocol had included a more salient means of establishing group identity, that involved recruiting six participants, randomly splitting them into two groups upon arrival, and running these groups simultaneously in separate rooms. However, in piloting this proved impractical due to

limitations with the available facilities and recruitment logistics. For study 2 it is recommended that these procedures be restored in order to promote a more salient 'group' identity, that will raise the baseline of affiliation across conditions.

Finally, fitting the HR monitors proved to be an extremely salient experience for many of the participants and a substantial portion of SER responses made at least some reference to heart rate or physiological responses. This is in certain respects unavoidable as participants will naturally try to puzzle out the purpose of any study they are involved with, and the fitting of HR monitors seems to provide a rather visceral clue. Efforts should be made in Study 2 to reduce the salience of this measure.

Study 2: 'Wailing Infant' Artificial Ritual V2 (Japan)

Study 2 was conducted in Hokkaido University in Japan and employed a randomly allocated, between subject design, with small groups of three participants completing a collective artificial ritual task, while affect was manipulated through a combination of auditory and visual stimulus. Following the ritual task, participants completed a questionnaire about their experience, various affiliation measures, an implicit association task (IAT) to examine positive associations with the group identity, and intra and inter-group allocation tasks using Tajfel matrices which were disguised as an independent study on memory and group cognition. All of these measures were also repeated again in a second session, after a one-week delay. Physiological arousal was measured during both sessions through the use of heart rate monitors.

Study 2 was originally intended as a cross cultural replication of Study 1 but due to the major limitations of Study 1 the methodology was revised significantly and the study

instead represented an independent conceptual test of the preliminary findings from the UK sample. The study was conducted in collaboration with an experienced Japanese social psychology laboratory and this resulted in additional changes being made in order to accommodate the distinctive cultural environment. The major changes to the protocol, and the motivations behind them, are detailed below and summarised in Table 3.12.

Protocol Revisions

Manipulation checks in study 1 suggested that two distinct conditions had been created but that the audio stimulus manipulation had proven relatively weak. This represented a fundamental concern for the validity of the study, as the ability to generate differential affective profiles between conditions is central to the study's purpose. Consequently, three substantial changes were made to the affective stimulus methods in order to try and strengthen affective manipulations.

First, the audio stimulus employed in the study was completely revised. There were several reasons for this decision but a central motivation was that although infant crying has been found to activate the sympathetic nervous system in adults (A. D. Murray, 1979; D Out et al., 2010; Philip Sanford Zeskind, 1985), not all crying is equal in its ability to induce dysphoric reactions in listeners (Boukydis & Burgess, 1982; A. D. Murray, 1985; Philip Sanford Zeskind & Lester, 1978; Philip Sanford Zeskind & Marshall, 1988). There is a significant body of research that focuses on responses to variations in infant crying (Owings & Zeifman, 2004; Zeifman, 2001), which has established that cries from more difficult infants, which have longer pauses within and between cry sounds and a higher fundamental frequency at peak intensity (Bates, Freeland, & Lounsbury, 1979; Lounsbury &

Bates, 1982; Philip Sanford Zeskind & Barr, 1997), result in greater discomfort and a more aversive response from listeners. There is also a substantial body of research into laughter which has similarly found that, although laughter is reliably found to serve as a contagious means of triggering positive affect in listeners (Provine, 1992, 1996a, 1996b), which can even increase pain tolerance (Dezecache & Dunbar, 2012), there are distinctive acoustic features of different types of laughter (Bachorowski, Smoski, & Owren, 2001; Szameitat et al., 2009) and these differentially elicit positive affect (Bachorowski & Owren, 2001).

Significantly for current purposes, studies have also found that young children produce at least four different types of laughter (Nwokah, Davies, Islam, Hsu, & Fogel, 1993) and that from amongst these it is the laughs with periodicity in vocal fold vibration (voiced) that most strongly promote positive affect amongst listeners (Bachorowski & Owren, 2001; Grammer & Eibl-Eibesfeldt, 1990; Hudenko & Magenheimer, 2011). Thus, the failure in Study 1 to elicit an appropriate euphoric response from the laughing infant stimulus during the stimuli pre-test could have been due to the wrong type of laughter being used.

In short, it was clear that greater precision was required for the infant audio stimulus. As such, Study 2 incorporated the use of stimuli from Oxford Vocal (OxVoc) Sounds database; a newly developed database of affective sounds, including infant vocalizations (Parsons, Young, Craske, Stein, & Kringelbach, 2014) with validated affective profiles. These stimuli were used to develop new audio tracks that were rigorously pretested in order to establish three distinctive affect profiles: dysphoric- infant crying (high arousal, negative valence), euphoric- infant laughing (high arousal, positive valence), and neutral- infant babbling (medium arousal, neutral valence). The euphoric condition was reinserted as it

necessary in order to establish whether any effects observed in the dysphoric condition are specific to negatively valenced arousal or apply to high arousal responses in general. Although these considerations could equally apply to justify the inclusion of the condition in Study 1 the priority during construction of that methodology had been on producing a streamlined design.

Second, changes in lighting conditions were added to further elicit the desired affective response. Darkness is a rare example of a universal cultural symbol which is found repeatedly across cultures to be symbolically linked with “evil, threat, and danger” (Schaller et al. 2003: 639). Such a symbolic association has been shown to have a range of psychological consequences, including: people in black uniforms behaving more harmfully (Frank & Gilovich, 1988); dark environments amplifying the acoustic startle effect (Grillon, Pellowski, Merikangas, & Davis, 1997; Grillon, 2008), even in virtual reality, (Mühlberger, Wieser, & Pauli, 2008); dark environments facilitating aggression (Page & Moss, 1976) and increasing prejudice against ethnic out-groups (Schaller et al. 2003; Schaller et al. 2003).

Conversely, a large and diverse research literature (Edwards & Torcellini, 2002) attests to the positive associations (P. Banerjee, Chatterjee, & Sinha, 2012; W.-B. Chiou & Cheng, 2013; Meier, Robinson, & Clore, 2004) and beneficial psychological effects of bright, natural lighting. For instance, increased access to windows and brighter lighting has been shown to be associated with: improved physiological, psychological, and behavioural health outcomes (Zadeh, Shepley, Williams, & Chung, 2014); increases in productivity and workplace satisfaction (Edwards & Torcellini, 2002; Haynes, 2008; Lottrup, Stigsdotter, Meilby, & Claudi, 2015); educational achievement (Hathaway, 1995; Slegers et al., 2012)

and tendencies towards ethical behavior (W.-B. Chiou & Cheng, 2013; Steidle & Werth, 2014). Moreover, a number of studies have demonstrated a preference for natural lighting over artificially-lit environments (Borisuit, Linhart, Scartezzini, & Munch, 2014; Edwards & Torcellini, 2002; Galasiu & Veitch, 2006; Gou, Lau, & Qian, 2013).

Consequently, to further distinguish the affective profiles between conditions, lighting conditions were varied in Study 2. Specifically, the dysphoric condition was conducted in ambient darkness, with lighting for visibility provided by a series of portable torches. Conversely, the euphoric condition was conducted under bright natural lighting, with all sessions performed at a time when there was bright daylight available. Finally, the control condition utilized moderate lighting supplied by artificial light sources, with any natural lighting from windows obscured. Promisingly, Richert, Whitehouse & Stewart (2005) also employed differential levels of lighting between conditions in two artificial ritual experiments and both studies reported significant differences between groups in the intensity of fright experienced.

Third, a wooden 'idol' with variable affective expressions was added to performances of the ritual sequence to increase the inferential richness of the event, invoke agentic reasoning, and increase ecological validity. One of the most well established findings in psychological research (Bruce & Young, 1998; Haxby, Hoffman, & Gobbini, 2000) and the cognitive science of religion (J. L. Barrett, 2004; Boyer, 1996, 2001; Guthrie, 1980, 1993) is the salience of faces to human cognition and the internal rumination that agentic cues prime through their enhanced inferential potential activating mental processes both of agency detection and 'theory of mind'. Studies have repeatedly shown that angry and fearful faces

prompt threat processing mechanisms generating negative affect (Lundqvist, Esteves, & Öhman, 2004; Mogg, Garner, & Bradley, 2007; Mogg, Philippot, & Bradley, 2004; Pinkham, Griffin, Baron, Sasson, & Gur, 2010) and happy faces promote positive affect and relieve stress (Becker, Anderson, Mortensen, Neufeld, & Neel, 2011; Hatfield et al., 1994; Papa & Bonanno, 2008; Surakka & Hietanen, 1998; Zajonc, 1985).

To make use of these implicit tendencies, the ritual sequence in Study 2 was modified to include a non-descript wooden 'idol' displaying a schematic face, based on illustrations provided in Maratos et al. (2008: 1343), that was angry (dysphoric condition), happy (euphoric condition), or neutral (control condition) (Fig.3.5).

Figure 3-5: Wooden idols with schematic faces



Furthermore, some actions were added in the ritual sequence in order to specifically draw attention to the idols, including some limited physical interaction in the final stage of the ritual performance. The enhanced inferential potential offered by the inclusion of the idol stimulus overall is intended to enhance the affective profile of the different conditions and to increase the ecological validity of the experiment, since agent representations are a recurrent feature of many popular religious rituals worldwide (Boyer, 1994, 2001; Guthrie, 1993).

Collectively, these three changes were designed to heighten dysphoric and euphoric arousal through the utilisation of stimuli corresponding to ‘ancient sources of harm’ which have been argued as the most effective at triggering evolved fear responses (LoBue & Rakison, 2013; Öhman & Mineka, 2001, 2003; Rakison & Derringer, 2008; Sander, Grafman, & Zalla, 2003). The manipulation between conditions remained variation in arousal and affect but in Study 2 was achieved through a combination of visual, audio, and lighting effects. A summary of the revised experimental conditions and the revised manipulations employed is provided in Table 3.11.

Table 3-10: Summary of Revised Experimental Conditions

Stimulus Category	Experimental Condition		
	Dysphoric	Euphoric	Neutral
Audio	Crying infant	Laughing infant	Gurgling Infant
Visual	Angry Idol	Happy Idol	Neutral Idol
Lighting	Ambient Darkness	Daylight	Artificial Light

The weakness of the experimental stimulus was not the only significant problem identified during Study 1. Instead, there were a series of concerns ranging from the overly

individualised nature of the collective ritual performance, to unintentional variations in the verbal demonstration provided. In total, there were seven areas that needed to be addressed, inclusive of the weak affective manipulations. Table X.XX provided after the methodology section summarises the amendments made to the protocol to address these issues and further detail of how the changes were implemented is provided in the experimental methodology section.

Hypotheses

H1: The highest levels of a) identity fusion and b) perceptions of relational bonds with other participants will be found in the dysphoric condition, followed by the euphoric condition, and lowest in the control condition (Whitehouse & Lanman 2014).

H2: The strongest in-group preference between participants will be found in the dysphoric condition (Whitehouse & Lanman 2014; Wiltermuth & Heath 2009; Xygalatas et al. 2013b).

H3: Higher levels of self-assessed positive affect will predict higher levels of a) identity fusion and b) perceptions of relational bonds with other participants (Chapter 2).

H4: Stronger physiological response to the artificial ritual will promote greater reflexivity and this effect will increase after a delay (Richert et al. 2005).

3.5 Methods

Ethics

The design for the study was based on the core methodology of Study 1 originally proposed by Harvey Whitehouse and Jon Lanman but was substantially revised by the thesis author, in collaboration with Masaki Yuki and his lab members at Hokkaido University. The current thesis author designed the revised protocol, produced all

experimental material, and supervised a team of graduate students from Masaki Yuki's lab, who conducted all of the experimental sessions⁵¹. Translations of material were produced independently by the study author, in consultation with native Japanese speakers, and were then revised in lab meetings at the Culture, Social Ecology, and Psychology lab at Hokkaido University.

The *Central University Research Ethics Committee* (CUREC) of Oxford University reviewed this research using human participants and approval was received for the experimental protocol (SAME/CUREC1A/12-14). All procedures conformed to the ethical codes prescribed by Oxford University. Additional approval was granted by the *Center for Experimental Research in Social Sciences* (CERSS) at Hokkaido University, who granted use of the Perception/Sensation Systems, the Worldwide Network, and Group Experiment laboratory rooms, under the supervision of Masaki Yuki.

A description of the study and what it involved, together with a full consent form was produced and provided to all participants prior to obtaining final agreement to participate. All participants were aware that they could end their participation, without penalty at any time, and were given the opportunity before, during, and after the experiment to ask questions. As the study involved deception, concerning the origin of the ritual sequence and the purpose of the performance, participants were provided with a detailed debriefing at the conclusion of the experiment and details to contact the experimenters if they had any further questions. The experimental manipulation was intended to generate dysphoric

⁵¹ This decision was taken as Masaki Yuki advised strongly against having any foreign experimenters present in the sessions because it would encourage stereotyped behaviour and introduce an outgroup prime. The thesis author was present in a separate room for all sessions and attended two early sessions due to equipment failure.

arousal and was anticipated to be unpleasant for participants; however, the negative effects were also designed to be temporary and non-enduring. All sessions were supervised by the current thesis author and conducted at Center for Experimental Research in Social Sciences (CERSS) at Hokkaido University.

Recruitment & Sample

108 male students ($M_{Age}=19.66$, $SD=1.418$), all of Japanese ethnicity and nationality, were recruited from a university participant pool public maintained by the Culture, Social Ecology, and Psychology lab at Hokkaido University. A single sex sample was selected due to recommendations from senior lab members in Japan that this would simplify recruitment, and that mixed sex experimental sessions risked reducing participant interaction, due to the low level of relational mobility in Japanese society (Yuki & Schug, 2012). The study was restricted to ages 18-35, but the limits were never invoked due to the university sample. As the study involved a potentially dysphoric experience those with a history of mental illness were also excluded. During recruitment participants were informed that the study required them to attend two sessions, with a one-week delay between each session and confirmed that they would be able to attend both sessions. For each six-person experimental session, seven participants were initially recruited in order to provide cover for cancellations. If all seven participants arrived, one participant was randomly selected and taken to an individual booth to complete a series of questionnaires, prepared by the members of the Culture, Social Ecology, and Psychology lab at Hokkaido

University⁵². If five participants arrived, a naïve undergraduate confederate was used to enable two sessions to run. And if fewer than five participants arrived, those who had shown up were given the option to reschedule for another session or to cancel their participation and receive partial compensation of 500 ¥. All participants in the full experiment received 1,500 ¥, comprised of a 1,000 ¥ participation payment and a 500 ¥ fixed bonus payment for the allocation tasks.

When a session began, the initial six participants were briefly altogether but after a short time were randomly allocated into two three-person groups. In this manner, each session involved two experimental conditions running simultaneously. There were 19 sessions in total, with 38 groups: 10 neutral, 14 dysphoric and 14 euphoric⁵³. Five groups involved confederates due to cancellations. All participants returned for the second session. The majority of the sample identified as non-religious (47.2%), 36.1% identified as Buddhist and there were small numbers of Christians (2.8%) and Shinto adherents (4.6%). The remaining 9.3% stated that they did not know.

Pre-Testing of Stimuli

For study 2, efforts were made to improve the quality and the affective profile of the infant vocalisation audio stimulus. In contrast to the audio tracks of study 1, which used open license sound effects, a validated set of non-acted infant affective sounds was obtained, through early access to the Oxford Vocal (OxVoc) Sounds database (Parsons et

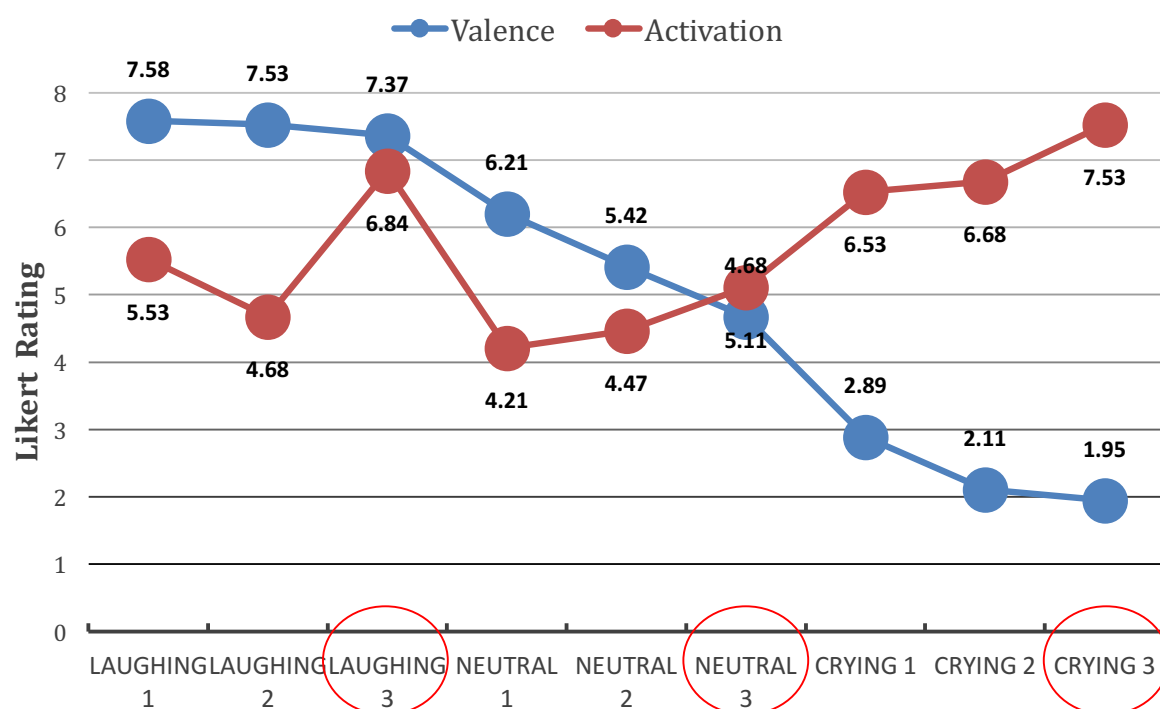
⁵² This was in accordance with established procedures at the Culture, Social Ecology, and Psychology lab. Participants selected for this received a partial payment of 1000 ¥ and the opportunity to rebook for another session.

⁵³ The unbalanced condition sizes were due to cancellations of two sessions that had both been randomly allocated 2 neutral groups.

al., 2014). This database contains short clips (1500 ms) of infant vocalisations, obtained from video recordings of infants filmed in their homes, with their primary caregiver, and is comprised of 25 neutral sounds, 18 crying sounds, and 18 laughing sounds. All clips display consistency in physical parameters and have been matched for total root-mean-square (RMS) amplitude at -25dBFS, which ensures consistency in overall intensity, “while retaining variations in amplitude envelope across time” (Parsons et al. 2014: 5). Additionally, all clips had been independently assessed by 34 raters for their level of arousal and valence and displayed the desired combination of high arousal-negative valence (crying sounds), low arousal-neutral valence (gurgling sounds), and high arousal-positive valence (laughing sounds).

However, when the OxVoc clips were serially combined into longer audio tracks, due to the repeated switching between different infant voices and short duration, the outcome proved disconcerting for all conditions. Consequently, the thesis author collected additional recordings from his nine-month old son during periods of intense crying, gurgling episodes, and euphoric laughter. These recordings were all collected in the same location, without any ambient background noise, and used the same recording equipment maintained at a constant distance. Using this additional audio material, interspersed with a selection of six clips from the OxVoc database, three audio tracks of approximately 120 seconds were created for each condition. All tracks were edited using Audacity software (version 1.3.4-Beta, <http://audacity.sourceforge.net/>), using the normalize function (set at -3dB) and manual editing of amplification to ensure a consistent maximum level of volume and soft fades between clips.

Figure 3-6: Affective profile of infant vocalisation stimulus



The nine tracks were then piloted on independent lab members ($N=19$: 12 Male, 7 Female, $M_{AGE}=25.75$, 16 Japanese nationals), who listened to each of the clips on headphones, and rated their affective response using semantic affect manikins (Bradley & Lang, 1994), with 9-point rating scales for affective valence and arousal (higher scores equate to more positive and more arousing responses). Following this, the mean valence and activation score for each track was calculated and these were plotted to identify the most appropriate track for each condition (Fig.3.6).

Following this, one way ANOVAs were conducted with both valence and activation ratings as outcomes and the individual tracks as between group factors. For valence, as expected, a main effect was found between tracks, $F(2,54)=55.643$, $p<.001$. Additionally, planned comparisons indicated that valence was significantly lower for the crying infant track than both the laughing and babbling tracks, $t(54)=9.165$, $p<.001$ and, crucially, the

laughing track was also found to have a significantly higher valence rating than the babbling track, $t(54)=5.223, p<.001$. For activation, again a significant main effect was found between tracks, $F(2,54)=13.203, p<.001$. Planned comparisons confirmed that there was significantly higher activation for the laughing and crying conditions, when compared against the babbling track, $t(54)=4.942, p<.001$, and no significant difference in activation between the laughing and crying tracks, $t(54)=1.409, p=.165$.

These results confirmed that the selected tracks were suitable in regards their affective characteristics. Specifically, valence ratings displayed a clear positive trend from dysphoric to neutral to euphoric stimulus and activation ratings indicated that both euphoric and dysphoric stimuli had higher arousal profiles than the neutral stimulus.

Experimental Procedure

Study 2 involved groups of six participants completing two experimental sessions separated by a one-week delay. During the first session, participants were randomly separated into two groups, comprised of three people, and completed a collective ritual together. This ritual group then stayed with their designated group for the remainder of both sessions. All sessions were organised around a cover story of their being two independent studies being conducted on the same day; the first exploring the meaning of recovered ritual instructions, organised by anthropologists from Oxford University, and the second examining memory, group cognition, and decision making, organised by the Psychology department at Hokkaido University. In reality, there was only one study that involved a collaboration between Oxford University and Hokkaido University. The cover

story was intended to disguise the connection between the experimental manipulations and the key outcome measures and thus reduce demand characteristics.

One assistant and three experimenters conducted each session, all of whom were Japanese nationals to avoid potential outgroup priming. All participants were informed that they would receive a minimum payment of 1500 ¥⁵⁴ (approx. £10) for their participation in the two sessions but that they might receive more, depending on certain decisions they made during the study

There were three experimental conditions (dysphoric, euphoric and control) that varied based on the affective character of environmental conditions visual, audio, and lighting conditions during the collective ritual performance (Table 3.11). In each session, two conditions were randomly selected and prepared in separate rooms, participants were randomly assigned to a condition based on a random selection performed during the experiment. The two experimenters who supervised the collective ritual performance were not blind to condition, as they prepared the room for the relevant condition and were present throughout the ritual task performance. However, the third experimenter, who collected the allocation and group identification tasks, ostensibly as part of a separate study, was kept blind to condition to reduce the potential for experimenter bias.

The experimental protocol involved following a six-stage process to ensure consistency between groups. A detailed description of each step and the measures included follows below.

⁵⁴ This amount was set according to conventions of the psychology department at Hokkaido University and was based on a total study duration of 1.5 hours (including both sessions).

Session 1

Step 1. Arrival

In a reception area, participants were welcomed individually by the assistant and provided with an anonymous ID and information sheets to read while waiting for other participants to arrive. All participants were given time to read the material and asked to indicate whether they still wished to participate. Recruitment material specified that the experiment required participants not have any prior relationship and that it was therefore unsuitable to sign up with friends for the same session. The experimenter checked this requirement with each participant and, when pre-existing relationships were discovered, informed the relevant experimenter, so that the individuals could be allocated different groups during the selection stage. The study information material provided an accurate summary of the contents of each session, however it was also designed to reinforce the cover story that two separate studies were being run simultaneously. As such, there were two information sheets and two consent forms for participants to sign. The first set referred to a “collaboration with an Anthropological research project at Oxford University, examining intuitions concerning archaeological artefacts” and included Oxford University logos, and the second set referred to a study on “the effects of memory on group cognition and decision making” being conducted by Hokkaido’s Social Psychology Lab, and included Hokkaido University logos. After consent forms were collected, the assistant reminded participants of the two study format, and told them they would now be starting the first study.

All participants were then taken to the group experiment room by the assistant, where they were greeted by the first experimenter. The experimenter then asked participants to take a seat at any one of six desks. Each desk was prepared with a laptop displaying the **pre-manipulation questionnaire**, a set of heart rate monitoring equipment, and a glass of water. Once all participants were seated, the experimenter introduced themselves, provided a brief summary of the experiment schedule, and then asked participants to enter their ID to the questionnaire.

Step 2. HR Fitting & Baseline collection

After entering their ID, the questionnaire then displayed an information sheet about the HR monitoring equipment, what it recorded, and how it should be fitted. The experimenter asked participants to read this and then provided a demonstration of how to correctly fit the chest transmitters. The desks in the group experiment room are designed to convert into private booths through the raising of sliding barriers and, as such, participants were able to fit the HR monitors in as much privacy as desired at their desks. The glass of water was provided so that the contact pads could be lubricated.

After all the transmitters were fitted, participants were shown how to use the HR wrist-receiver, and asked to enter their baseline measurements in the questionnaire. When any technical problems were encountered the experimenter provided assistance. Providing the HR equipment at this early stage was intended to allow for some degree of habituation before the actual ritual performance. The experimenter also emphasised at this point that, if the HR monitor stopped working at any part of the experiment, it was not a significant issue, and that the participant should continue as normal. This instruction was intended to

reduce the importance attached to the HR monitors, based on the importance attached to the monitors by some participants during study 1.

After all monitors were fitted and baseline measures completed participants were then asked to complete the remainder of the **pre-manipulation questionnaire**, which contained the following measures:

- (1) Demographic questions (age, sex, nationality, ethnicity, religious affiliation, fluency).
- (2) The 10-item International Positive and Negative Affect Schedule (I-PANAS-SF) (Thompson, 2007) to measure baseline mood, using initial translations from Clark & Watson (1989).
- (3) The single-item Affect Grid (J. A. Russell et al., 1989) used to measure baseline affect and activation.

Step 3. Cover Story and Group Assignment

After completing the **pre-manipulation questionnaire**, participants were then provided with an elaborated account of the cover story by the experimenter. Specifically, the experimenter explained the following:

A team of anthropologists and archaeologists from Oxford University have recently excavated two previously unknown and closely located archaeological sites. At the sites, a series of instructions for two previously unknown rituals, preserved on dried animal skins, have been found. Although they know which groups inhabited the sites and have been able to reconstruct the full set of instructions concerning how to perform the rituals, they have very little information about why people performed the rituals, or what the various actions described may have symbolised.

One way anthropologists try to reconstruct the purpose of ancient rituals is to examine people's intuitions about them after they have performed them. Consequently, Hokkaido's Psychology department, in collaboration with Oxford University's School of Anthropology and Museum Ethnography, will be asking participants to perform a reconstruction of the rituals described in the recovered instructions. The reconstructed rituals will be performed in small groups of three people, as specified by the recovered instructions, and afterwards participants will be asked to answer a series of questions on their thoughts about what the ritual may have meant and how they responded to it.

To help increase realism (within the confines of the lab space), certain elements (e.g., sound effects & lighting conditions) will be used to partially simulate the environment in which this rituals are likely to have originally been performed. However, please note that some specific cultural details have been removed to avoid any potential biasing of responses. The setting and the items used are not the original artefacts but rather approximations provided by anthropologists at the University of Oxford. Participants will be asked for their immediate impressions, and their impressions after having some time to reflect on the experience, as such the study will take place over 2 sessions: one week apart.

The experimenter then emphasized that the study was examining two different sets of ritual instructions, and that that these were associated with two different tribal groups; the *Zanshii* (ザンシイ) and the *Kokair* (コカイー).⁵⁵ These 'tribal' names were selected based on similar groups used in Pinter & Greenwald (2011) and following pretesting of a large list of

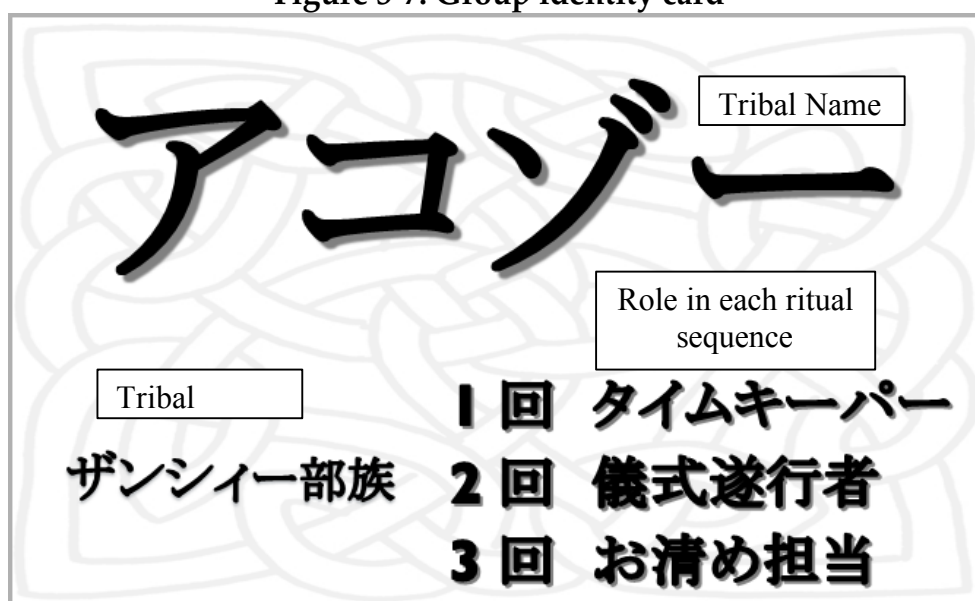
⁵⁵ The characters used to spell these names were katakana, which is a script in Japanese associated with foreign words.

alternatives with the members of the Culture, Social Ecology, and Psychology lab. The chosen group names were judged to be memorable, not identifiable with a known culture, and to have no problematic associations in Japanese. These group names were also required for the IAT task and enabled a simple heuristic to be applied to distinguish members of both groups based on the methodology described by Pinter & Greenwald (2011).

To select which ritual the participants would perform, they were asked to draw a laminated card (Fig.3.7) bearing a 'tribal' name from a choice of six (3 Zanshii & 3 Kokair) cards contained in an opaque box. The order of selection by participants was determined by seating which was freely chosen by participants when they arrived in the room. The intent with this procedure was to provide a salient group identity that was arbitrarily determined, but also selected freely by the participant. The only exception to this procedure was when the assistant had identified participants with previous relationships and, in these cases, if the known individuals randomly selected the same group, their cards were swapped with another participant in the opposing group.

In addition to the tribal group, each card also contained an individual tribal name, prominently displayed at the top of the card (Fig.3.7). These names were used to determine the order in which participants performed the main ritual sequence, which was also indicated on the card, and to provide a further personalised connection with the group identity.

Figure 3-7: Group identity card



Note: English notations added for clarity

Additionally, the names selected employed a simple heuristic that connected them to the relevant group, specifically all of the names allocated contained within them the first character of the group name. For the Zanshii group, this meant all the names contained *Za* (ザ), whereas for the Kokair group, all the names contained *Ko* (コ). The names designated in each group were matched by syllable length and position of the group relevant character (Table 3.13). Participants were asked to note their names and were told that their relevance and the meaning of the roles listed, would be explained shortly.

Table 3-11: Tribal Names & Roles in first ritual task

Role in 1 st Ritual Sequence	Zanshii Ritual Names	Kokair Ritual Names
Main Performer	<u>Z</u> aritto / ザリット	<u>K</u> orakka / コラッカ
Timekeeper	A <u>z</u> ar / アザアー	A <u>k</u> ozo / アコゾー
Purifier	Re <u>z</u> a / レザー	Ki <u>k</u> o / キコー

Following the selection task, the two groups were separated. One group remained in the group experiment room, while the other was led to the physical experiment room by the assistant. The name of the group that left the room was counterbalanced between sessions.

Step 3. Collective Ritual Performance

Before the session, both the group experiment room and the physical room were prepared for the ritual performances with the necessary apparatus and stimulus equipment laid out identically in both rooms. The same surround speakers were used in both rooms and were positioned approx. 1.5 metres from the ritual apparatus with volume settings kept consistent. An A2 poster, which outlined all of the steps in the ritual sequence, was also hung from the wall at kneeling height, at a distance of approximately half a metre from the ritual apparatus. As both rooms were used for all three conditions, a series of portable lights had also been positioned around the room to illuminate the poster and the ritual apparatus, during the low lighting of the dysphoric condition. To shield these areas from being viewed in advance partition curtains were used. In the group experiment room, a partition running across the length of the room disguised the ritual apparatus and in the physical room, mobile partitions cordoned off the apparatus from a reception area, comprised of three desks with laptops and chairs.

Before being led to these areas, the participants in both rooms were asked to put on headphones and watch a 5-minute instructional video to show them how to perform the ritual correctly. The video was framed as being prepared by the anthropologists at Oxford for use with international collaborators and featured three caucasian graduate students

(including the thesis author), performing the ritual sequence in silence, with instructional subtitles in Japanese (Fig. 3.8).

Figure 3-8: Still image from instructional video



Before the actual ritual sequence was demonstrated, the video began by reminding participants that they would be asked to replicate the procedure demonstrated and should thus pay careful attention to what is shown. On this first screen, participants were also told that they should:

1. Not talk, or communicate verbally with all participants, during the ritual performance.
2. Treat the ritual performance seriously and respectfully.
3. Try to recall as much of the sequence as possible from memory and only refer to the written instructions when necessary.
4. Note that the main ritual sequence will be repeated three times, and that each participant will perform all of the ritual roles once during the performance.
5. Note that the colour of the powder is different during each repetition. Yellow in the 1st sequence, Red in the 2nd sequence, and Blue in the 3rd sequence.

The next screen indicated to participants that some environmental conditions, including background sounds and ambient lighting would be “altered to more accurately stimulate the original context” of the ritual during their performance but that the content otherwise would remain the same. The revised version of the artificial ritual involved three interactive roles: a central ritual performer, a purifier (cleaner) and a timekeeper, and the second screen emphasised that each role would be performed once by all participants. The addition of the timekeeper and purifier roles was primarily intended to encourage those not performing the central ritual sequence to remain attentive and to increase their sense of involvement across the entire ritual performance.

Following the second instruction screen, the actual demonstration video began with the full ritual sequence being broken down into four stages that were repeated three times: 1. Preparation, 2. Central Performance, 3. Purification, 4. Changing Roles, before the final stage: 5. Conclusion of the Ritual. The full ritual sequence including all three repetitions was demonstrated, but various segments were sped up to reduce the duration of the video. During each stage, subtitles provided a summary of what was being performed and by whom. To see a full breakdown of all instructions provided please consult Appendix III but below a list of necessary materials is provided and a shortened summary is provided.

Figure 3-9: Ritual Material initial layout (euphoric condition)



The physical materials for the ritual sequence comprised of the following:

- (1) A small black table with legs collapsed (used as the platform for the main ritual sequence).
- (2) 3 custom wooden idols, with interchangeable heads displaying positive, neutral, and negative schematic faces.
- (3) 1 x large wooden bowl, wrapped with twine, filled 1/3rd with water and placed on top of a small white pot holder.
- (4) 1 x medium bowl, used to mix materials during the main ritual sequence.
- (5) 1 x small bowl, used to transfer the
- (6) 5 x small glass jars with 2 x spoons that contained i) odourless red, blue and yellow dyes, and ii) assorted birdfeed seeds.
- (7) A weaved basket that contained plant material gathered from nearby shrubs.

- (8) 2 x weaved holders, for holding a small wooden stick and the plant material.
- (9) A black granite mortar and pestle for grinding the seeds.
- (10) 3 x small square wooden
- (11) 3 x leather necklaces
- (12) 1 x large plastic bowl (for refilling water)
- (13) 1 x clear plastic bowl (for holding towels)
- (14) 2 x towels for cleaning ritual apparatus between performances
- (15) A small circular metal bell and a metal striker
- (16) 2 x wooden clapping boards
- (17) A bamboo mat
- (18) 3 x weaved wooden circular cushions

Heart rate measures (avg. and max) were collected using 6 x coded Polar FT2 heart rate monitor watches, with matched T31 chest band transmitters.

The central ritual sequence performed in Stage 2 remained similar to that used in the previous artificial ritual experiment but a number of changes were made (Table 3.12). First, instructions were added to “stare intently at the face of the wooden idol” during the stirring segments. This was done to increase awareness of the relevant schematic face. Second, participants used a stick instead of their hand to stir the mixture, this was done to avoid participants from staining their hands in the dye. Third, the portion of the ritual that involved waiting for a water timer was also removed, as this was found during piloting to extend the duration of the ritual to an unacceptable length. Fourth, participants were requested to blow on the plant material when transferring to the mortar bowl. This step

was added to try and provide some sense of reverence for the materials, this is also the reason why the stirring stick was returned to a holder and why some implements were elevated on holders, such as the collective bowl. Fifth, the coloured mixture produced during the performance was not discarded but instead was transferred into a collective bowl, that was later used in Stage 5, after all three repetitions had been completed. This was intended to increase the sense that the ritual was a collective activity, as the mixture in the collective bowl continued to change colour after each repetition due to the different coloured dyes.

The addition of Stages 1 & 3 was due to the preparation and cleaning activities, that in study 1 were performed by the experimenter, being repackaged into tasks for the 'purifier' to reduce external interruptions. The timekeeper now also initiated each stage by ringing a bell, and marked the changing of roles and the conclusion of the ritual by the clapping of wooden blocks together. These tasks were added in order to provide an orientating structure for the participants and to increase the feeling that the ritual was a collective activity. This was also part of the motivation for adding Stage 5, in which all participants together lifted the idol, placed it in the collective mixture, and took turns using the stick to sprinkle drops of the mixture on the idols head. Another significant motivation for adding this stage was to to increase interaction with the idol to invoke a greater sense of agentic involvement and potentially increase reflection about the purpose of the ritual.

Table 3-12: Instructions for Main Ritual Sequence (Stage 2)

Step	Action
1	Pour the powder into the water in the large wooden bowl on the altar.
2	Pick up the stick and, holding it in your right hand, stir the mixture in the bowl, 25 times, going clockwise. Stare at the face of the wooden idol intently while doing so. When finished return the stick to its holder.
3	Grind the seeds carefully in the mortar, using the pestle. Use your hand to prevent the seeds from spilling and pour the ground seeds into the mixture in the large wooden bowl.
4	Pick up the plant material, raise it slowly to your face then blow gently on it and place it into the mortar.
5	Grind the plant material thoroughly and then transfer the ground plants into the mixture in the large wooden bowl.
6	Pick up the stick, and holding it in your right hand, stir the water 25 times in an anti-clockwise direction. Stare at the face of the wooden idol intently while doing so. When finished return the stick to its holder.
7	Transfer the mixture prepared in the large wooden bowl to the collective bowl, wrapped with string and located in the centre of the mat.

Once the video of the ritual performance ended a final screen appeared that informed participants that mistakes were common during replications and that they need not be overly concerned if they make mistakes but rather proceed on as best they can. Participants were also reminded that the experimenter would not be involved with the performance in any way and that therefore participants would have to decide when to begin new stages, change roles, and so on. Finally, participants were warned that various environmental effects, in term of lighting and background noise, would be added based on the instructions provided by the anthropologists.

Participants in each group were then led by the experimenter through the partitions to the prepared ritual staging area. Participants were asked to sit on the floor in a triangular formation, taking the position indicated by the cards they had selected (purifier on the left, timekeeper on the right, and performer in the centre) (Fig.3.10). The video had instructed participants to change roles in Stage 4 by rotating positions clockwise and the experimenter reminded participants of this instruction. After all participants were seated, the experimenter explained that the ritual performance would begin when all environmental conditions, including the background audio track were ready. Participants were given the opportunity to ask any final questions before the performance began and were reminded that instructions were visible for all stages on the poster hanging from the wall, if necessary. Finally, just before the audio file was activated participants were requested to activate their HR monitors.

Figure 3-10: Positions for ritual performance



Participants then completed the collective ritual accompanied by the euphoric, dysphoric, or neutral audio track, relevant lighting conditions (natural light, room lighting,

or ambient darkness), and the appropriate schematic face displayed on the wooden idol. The experimenter observed the performance from a position partly obscured from the participant's point of view and recorded any errors made in the protocol but did not intervene unless this proved unavoidable. The full ritual performance took approximately 15 minutes to complete. When the performance was finished environmental conditions were returned to normal and participants were requested to remain seated, and to record on a sheet provided the three HR measures provided by their wrist monitor: i) their current heart rate, ii) average heart rate during the ritual, and iii) maximum heart rate during the ritual. The experimenter then led all three participants back to the individual tables where they had watched the instructional video.

Step 4. Post-Ritual Measures (IAT) & Dismissal from 'Study 1'

Once seated at the tables, participants were directed to begin the post-ritual questionnaire completed on a laptop. This questionnaire began with an Implicit Association Task (IAT) run on Inquisit V4 Software that was designed to assess the degree to which participants held implicit positive associations about the tribal group identity they had been allocated to. In essence an IAT task involves participants being asked to quickly classify different concepts into the correct category by pushing one of two keys which correspond to categories that are presented on the left and right hand side of the screen. The categorisation is completed in blocks and between blocks the position of the categories change. Blocks can involve single categories on each side or a pair of categories. The speed with which the task is completed and the errors made are then used to calculate a D score which indicates how closely respondents rated selected types of concepts. Due to cognitive

demands an IAT is believed to provide an implicit rather than a reflective measure of association. In this task the paired target concepts used were Zanshii tribe vs. Kokair tribe and Positive vs. Negative and these were presented in two blocks of 40 trials, for a total of 80 trials. The task was introduced to participants with the following statements:

We are interested in how completing the ritual task affects your ability to correctly identify names associated with the tribe whose ritual you just performed. As such, you will be shortly provided with instructions and asked to perform a categorizing task of tribal names/titles.

Then the heuristic, already briefly introduced during group allocation, was restated:

There is an easy way to identify members of the Zanshii (ザンシー) tribe as all of their names contain the letter ZA (ザ). Similarly, all the members of the Kokair (コカアイ) tribe have the letter KO (コ) contained somewhere in their names. Please try and concentrate on the names provided and remember the simple categorizing rule.

Participants were then given 60 seconds to familiarise themselves with a list of sixteen names (eight for each tribal group) that were used as targets in the IAT task (Table 3.13). This time limit was not intended to enable participants to memorise the names but simply to increase familiarity with the heuristic and prepare them for the format of the targets used in the IAT.

Following this, all participants individually completed a Brief Implicit Association Test (BIAT) (Sriram & Greenwald, 2009) using Inquisit V4 presentation software. The BIAT had four categories, combining a concept classification (Zanshii vs. Kokiar) with an attribute classification representing positive vs. negative valence (Good vs. Bad). The BIAT instructs respondents to focus on just two of the four IAT's category–response mappings in each

combined task. Subjects are shown two category labels with their exemplars at the beginning of the task and are instructed to “keep them in mind”, to respond to items from these two categories with a “focal” response key, and to respond to any other stimuli with an alternative “non-focal” response key. Two blocks of fewer than 80 presentation trials are then conducted, substantially reducing the necessary administration time as compared with the standard 5-step IAT procedure which involve seven blocks of trials amounting to 180 trials.

Table 3-13: IAT targets

Kokair Names	Zanshii Names	Good 良い	Bad 悪い
<u>B</u> ekor / ベ <u>コ</u> アー	<u>B</u> izaten / ビ <u>ザ</u> テン	Marvelous/ 素晴らしい	Tragic/ 悲愴な
Nor <u>z</u> kos / ノル <u>コ</u> ス	<u>M</u> er <u>z</u> as / メル <u>ザ</u> ス	Superb/見事	Horrible / 恐ろしい"
<u>K</u> omer / <u>コ</u> メール	<u>Z</u> alar / <u>ザ</u> ラル	Pleasure/喜び	Agony/苦しみ"
<u>E</u> kok / エ <u>コ</u> ク	<u>A</u> zaj / ア <u>ザ</u> ージ	Beautiful/きれい	Painful/痛い
<u>R</u> eko / レ <u>コ</u>	<u>K</u> aza / カ <u>ザ</u>	Joyful/楽しい	Terrible/ひどい
<u>K</u> oni / <u>コ</u> ニ	<u>Z</u> ani / <u>ザ</u> ニー	Glorious/ すばらしい	Awful/いやな"
<u>A</u> kol / ア <u>コ</u> ル	<u>I</u> zari / イ <u>ザ</u> リ	Lovely/ すばらしい	"Humiliate/ 恥ずかしい"
<u>K</u> olkiz / <u>コ</u> ルキズ	<u>Z</u> andix / <u>ザ</u> ンディックス	Wonderful/豪華	Nasty/下劣な

When the BIAT was completed, participants were instructed to continue with the post-ritual questionnaire, which included the following six measures:

- (1) The 10-item International Positive and Negative Affect Schedule (I-PANAS-SF) (Thompson, 2007) to measure post ritual mood.

- (2) The single-item Affect Grid (J. A. Russell et al., 1989) to measure post ritual affect and activation.
- (3) The open ended SER measure, described in Study 1.
- (4) The four relational bonds items, described in Study 1.
- (5) Empathetic and Distress response to infant sounds, described in Study 1 (Batson et al., 1983).

The questionnaire requested participants to signal to the experimenter after they had completed all questions and then to wait in their seat for further instructions. Once all participants had signalled their completion the experimenter thanked them for their participation, asked them to remove their HR monitoring equipment, and provided them with an appointment card for the second session. This card included the time and date of the second session, specified the group name, and listed the three tribal names associated with the group. The appointment card served an instrumental purpose as a reminder but it was also intended to provide a visible cue of the experience, potentially increasing reflection during the 7-day gap. Asking participants to remove the HR monitors was further intended to serve a symbolic function and provide support for the cover story concerning the existence of two independent studies.

Step 5. 'Study 2' (Fusion & Tajfel Allocation Tasks)

Once all HR monitors were removed, the assistant escorted participants to the entrance of the individual booth room. Here, they were greeted by the third experimenter, who

welcomed them to start of the second study⁵⁶ and provided a brief orientation about the overall structure of the study. They were informed that the study would ask questions about groups they had belonged to for different lengths of time, including the group from the previous experiment, and that they would then complete a series of memory tasks, through which they could a bonus based on their performance. Participants were then sent to individual cubicles in which a PC displayed the allocation questionnaire.

The allocation questionnaire began by reiterating the structure of the study outlined by the experimenter using the following instructions:

This study will be conducted over two weeks and consists of two parts:

- *In the first part of the study: you will be asked some questions about your feelings towards different groups to which you belong. Some of the questions will refer to long term groups and some will refer to short term groups including the groups you were allocated to during the previous experiment.*

Specifically, in this section, participants completed brief group identification and identity fusion scales with three separate target groups: their experimental group (Kokair or Zanshii), their university (Hokkaido), and their country (Japan). The identification measure was comprised of the single-item measure of social identification (Postmes et al., 2013) and three representative items from the in-group identification scale (Leach et al.,

⁵⁶ Initially the protocol required all six participants from both groups to arrive before Study 2 was introduced but this was abandoned after the first 2 sessions due to significant delays in finishing times between groups. Instead, for all the remaining sessions, each individual group was allowed to begin study 2 as soon as they arrived at the individual booth room. This decision was taken as participants completed all of the procedures in this section of the experiment using PCs arranged in isolated rooms so there was already very little interaction between the participants.

2008)⁵⁷. This particular combination was chosen due to the wide usage of these two scales in recent research, a desire for brevity, and recommendations provided by Tom Postmes, via personal correspondence. The items included were:

1. *I identify with my [group]. (Identification-Postmes et al. 2013)*
2. *I feel committed to my [group]. (Solidarity- Leach et al. 2008)*
3. *I am glad to be a member of my [group]. (Satisfaction- Leach et al. 2008)*
4. *Being a member of my group is an important part of how I see myself. (Centrality- Leach et al. 2008)*

The identity fusion scale was also reduced, in consultation with Bill Swann, to a 4-item version of the verbal fusion scale (Gómez, Brooks, et al., 2011), which included the following items:

1. *I am one with [my group].*
2. *I'll do for [my group] more than any of the other members would do.*
3. *I have a deep emotional bond with [my group].*
4. *I make [my group] strong.*

The second part of the study was also summarised, as follows:

- *In the second part of the study: most participants will complete a series of memory tasks which involve memorizing a list of items or details from short stories and then answering questions related to these. You can earn bonus payments based on your performance in these tasks.*

Following this the study summary also introduced a surprise alternative task that one participant from the group would be randomly selected to perform instead of the memorisation task. This task was described as “a series of resource allocation tasks” and it

⁵⁷ This combination was selected over the Mael & Ashforth scale (1992) used in chapter 2, as during pretesting Japanese translations several of the items were found to be incompatible with the experimental context, such as, references to criticism in the media. Additionally,

was explained that whoever was selected for this “cannot earn any bonus payments directly but will instead be paid a fixed bonus, which does not change regardless of any of the choices made”. In actual fact, all participants were ‘selected’ to complete the allocation tasks. This ‘lottery task’ procedure was based on a method previously employed by Yuki & Yokota (2009) and was designed to avoid confounding motivations in allocation choices based on self-interest or the maintenance of reciprocal exchange with other ingroup members (Jin et al., 1996; Karp et al., 1993). The procedure avoided these confounds by making the participants aware that their payment was already fixed (therefore eliminating self-interest), and clarifying that no other participants in their group would be making allocations (eliminating any potential for future reciprocity).

All participants were informed that they had been selected to complete the allocation task after completing the identification and fusion measures. The instructions explained that they would perform this task in both sessions and that they “should not reveal [their] role to any of the other participants in the study or discuss what the task involves” until after the experiment was complete. It was also restated that their bonus payment amount would be set at a fixed amount, regardless of their choices, and that all the choices they made would be kept anonymous. To increase motivation participants were informed that their decisions would “have real consequences on the bonus amounts paid to other participants” and should thus be selected carefully, but that there was “no right or wrong answer”.

On the next screen, participants were provided with instructions explaining the structure of the allocation tasks, which were based on Tajfel Matrices (Billig & Tajfel, 1973;

Tajfel et al., 1971a; Tajfel, 1970). Tajfel Matrices ask participants to choose from one of thirteen payment options that award amounts to two paired targets. Participants make repeated choices in allocations from vertically paired choices. There are various types of matrices and in some the amounts awarded in each payment option are linked, meaning that as payments to one target increase, payments to the other target decrease (Matrix Type A) (Fig 3.11).

Figure 3-11: Example Tajfel Matrices

Matrix A

Please select a bonus allocation for both participants indicated from any of the vertical pairs. (Amounts in ¥)

INGROUP MEMBER	190	180	170	160	150	140	130	120	110	100	90	80	70
OUTGROUP MEMBER	10	30	50	70	90	110	130	150	170	190	210	230	250

Matrix B

Please select a bonus allocation for both participants indicated from any of the vertical pairs. (Amounts in ¥)

INGROUP MEMBER	70	80	90	100	110	120	130	140	150	160	170	180	190
OUTGROUP MEMBER	10	30	50	70	90	110	130	150	170	190	210	230	250

Alternatively, there are matrices (Matrix Type B) (Fig. 3.11) that contrast maximum difference in allocations with maximum joint profit, meaning that participants must choose between awarding a target a larger overall payment or a payment which is smaller in absolute value but larger relative to that awarded to the opposing target.

Multiple matrices are presented and the position of the targets is usually counterbalanced, for example by having ingroup members as the top target and then later using the same matrix but placing the outgroup member as the top target. By analysing decisions made over a combination of matrices 'pull scores' (Bourhis et al., 1994) can be

calculated that can be used to distinguish between various motivations. For instance, it is possible to distinguish between the strength of desire to award the maximum profit for a specific target from a desire to award maximum joint profit to both targets, or alternatively the relative *pull* for generating maximum difference in allocations in favour of a specific target. Typically, the matrices have been used with ingroup vs. outgroup targets and in this manner they can be used to assess intergroup discrimination, as this is reflected by a stronger pull towards maximizing difference in favour of the ingroup, as opposed to basic ingroup bias, which is instead demonstrated by a stronger *pull* towards allocations that maximize ingroup profit.

In the current study, participants were presented with the following three paired targets:

- 1) A member of their ritual ingroup vs. A member of the other ritual outgroup
- 2) A participant from their university vs. A participant from another university
- 3) A Japanese participant vs. A non-Japanese participant

For each pair of targets, participants were presented with four different matrices, two based on Type A matrices that contrast maximum ingroup profit and maximum difference with maximum joint profit and two based on Type B matrices that contrast maximum difference with maximum joint profit. Each of the matrices was presented with the target on the top row counterbalanced to avoid the potential confound of choices being made based on the presentation of targets.

The instructions provided to participants were introduced as follows:

You will now complete an allocation task which involves allocating bonus payments between two group members.

In each of the following rounds you will be presented with a table containing two rows with 13 columns (see below) containing numbers which signify bonus payments (in ¥) to be paid to other participants of the study. Although you will know the group the other participants belong to you will not know the specific identities of the individuals you are allocating bonuses too as every participant has been allocated an anonymous code number.

Example:

表中の数字はボーナスの金額を表しています。

(¥)

グループ X のメンバーの74番

グループ Y のメンバーの66番

190	180	170	160	150	140	130	120	110	100	90	80	70
10	30	50	70	90	110	130	150	170	190	210	230	250

The two rows indicate the amounts to be allocated to two different participants and you can select any of the 13 bonus options available. However, you must select the payment in the same column for both participants. For instance, if you chose to allocate 100 ¥ to Member 74 of Group X then you would also automatically chose to allocate 190 ¥ to Member 66 in Group Y. To select the option you desire you should click once anywhere inside either box within the chosen column.

To increase involvement participants were also informed that:

From the choices you make a random sample will be selected and will have direct consequences to the bonus amounts paid to other participants at the end of the study. There is no way to tell which of the questions will have real consequences, so please treat each round seriously and make the choice that you consider best.

Participants were then given two practice allocation tasks and tested to see that they correctly understood how the allocations worked. If they failed these tests they were provided with additional instructions and asked to contact an experimenter, by pressing a button located on the wall, if they required additional assistance. Once the practice tasks were completed successfully participants were able to proceed to the actual allocation tasks. After these were completed, a final question asked participants to select their preference for which target should receive the actual bonus payment. But there was no indication provided that this would influence the decision.

The last screen on the questionnaire, in accordance with standard protocol for Japanese questionnaires contained demographic questions (*age, sex, nationality, university year, living arrangements, household wealth*). These were included only to enhance the impression that the two studies were independent.

Once the demographic section was completed the questionnaire thanked participants for their participation and asked them to call the experimenter, via a button on the wall. The experimenter arrived, reminded participants about the second session, and then they were escorted to exit the booth room.

Step 6. Session 2

A week later participants returned for the second session. Appointment cards were collected upon arrival, and all participants were checked to ensure that they could recall which tribal group and tribal name they had been allocated. If anyone was uncertain they were reminded before proceeding to the relevant experimental room.

In this session, Steps 5 & 6 were repeated in their entirety (including the IAT) using the same rooms as session 1, with instructions amended to reflect that the tasks were being repeated for a second time. The actual ritual performance was not repeated, nor were participants required to wear heart rate monitors. The only changes to the questionnaires were that the post-ritual questionnaire finished by asking participants to indicate if they had any previous relationships with any of the other participants in the experiment. For the allocation questionnaire, after completing the group identification and identity fusion measures, participants were reminded that they had been selected in the previous session to complete the allocation tasks and would now be asked to repeat the same role. The instructions also specified to participants that they were not required to make the same choices as in the previous session and could choose their responses freely. The instructions for the Tajfel matrices were then repeated, along with the practice tests, and then participants made allocations using the same matrices and the same three paired targets from session 1.

Finally, participants were asked two questions concerning whether they found any aspect of the experiment difficult to understand and what they believed the purpose of the two studies was. Once completed participants were provided with a detailed debriefing sheet that explained the true nature of the study, the reason for the deceptions, and were given the opportunity to discuss any remaining questions with the experimenter. Once satisfied, participants were issued with their payments for participating and asked not to disclose the contents of the study to others until after all sessions had been completed in a month's time.

Table 3-14: Summary of Methodological Improvements

Limitation identified in Study 1	Amendments made for Study 2
1. Weak affective manipulations	- Revised audio stimuli created incorporating sounds with validated affect profiles. - Ambient lighting conditions manipulated. - Affect appropriate schematic faces, added to wooden idols to provide visual affective cue.
2. Ritual task too individualised	- New interactive roles added so that each ritual sequence involved all participants. - Different coloured dyes used in the ritual sequence by each performer changed the colour of the water in a shared collection bowl. - Collective action added to conclude the ritual performance.
3. Purpose of the study too evident	- Allocation tasks disguised as part of an independent study. - International collaboration emphasised in experimental materials to enhance credibility of cover story.
4. Variation in ritual demonstrations	- Ritual instructions provided individually through a standardised instructional video.
5. Potential influence of researcher bias/demand characteristics	- Additional experimenters used to enable main experimenters to remain blind to condition. - Interactions with experimenters were reduced substantially; all questionnaires, including allocation tasks, completed in individual booths. - Experimenter no longer interrupts ritual performance.
6. Limitations with DVs	- All core DVs were collected immediately following ritual performance and again after a week delay. - Tajfel matrices used for finer grained analysis of prosocial actions and examine intergroup discrimination. - Verbal fusion and identification measures added for various target groups. - Implicit ingroup preference (IAT) measure added.
7. Lack of clear group identity	- Participants randomly allocated into two named ritual groups. - Cover story used to create clear ritual 'outgroup'. - Tribal names provided and used to identify ritual 'roles' - IAT task uses tribal names and in-group/out-group members. - Appointment card for session 2 contains reminder of ritual group name and ritual roles.

3.6 Results

Data Preparation

Prior to analysis, the demographic composition of the three conditions was examined for any significant differences in age, household wealth, religious background but no significant differences were detected, using one way ANOVAs.

Relational bonds

Based on the results reported in study 1, the four items intended to measure relational bonds between participants were combined into a single factor scale which was found to have a moderate to excellent reliability across both sessions (S1, $\alpha=.796$ and S2, $\alpha=.890$).

Group Identification

The 4-item combined group identification scale was tested for reliability with all targets and was found to have excellent reliability when combined into a single factor scale (Ritual Group- S1, $\alpha=.926$ and S2, $\alpha=.952$)

Identity Fusion

The 4-item verbal identity fusion was tested for reliability with all targets and was found to have excellent reliability when combined into a single factor scale (Ritual Group- S1, $\alpha=.939$ and S2, $\alpha=.901$).

Implicit Association Task

Measures of association strength based on BIATs were computed using the D measure, an effect-size-like measure which ranges in value from -2 to $+2$ (Greenwald, Nosek, & Banaji, 2003). Due to the focal attribute classification being allocated to positive valence, higher scores equate to a more positive association with the two concept categories: the

ritual in-group and ritual outgroup. D is calculated “as the difference between mean latencies of the two BIAT blocks divided by the inclusive (not pooled) standard deviation of latencies in the two blocks” (Sriram & Greenwald 2009: 286).

Positive / Negative Affect

Positive and negative affect measures were collected using two measures at three time periods; upon arrival to establish a baseline, after the ritual performance to provide a post-manipulation measure, and after a 1-week delay to enable examinations of consistency.

The first measure was a single item affect grid (J. A. Russell et al., 1989), which enables affect and activation to be classified based on the position selected on a four quadrant grid. Responses were coded for affect valence and activation using values ranging from -9 to 9.

Responses to the 10-item I-PANAS-SF measure were analysed according to established protocols (Thompson, 2007); aggregating the responses from the 5-positive items ($\alpha=.887$) and the 5-negative items ($\alpha=.851$) into independent positive and negative affect scores.

Cooperation / Ingroup Preference

Using responses to the four paired Tajfel Matrices first, the total amount selected to be donated to various ingroup and outgroup targets were calculated. Second, ‘pull scores’ were calculated in line with methods specified in Bourhis, Sachdev, & Gagnon (1994) that represented:

- 1) Pull of ingroup favouritism against maximum joint profit (Matrix A)
- 2) Pull of maximum joint profit against favouritism (Matrix A)
- 3) Pull of maximum difference (intergroup discrimination) against maximum joint profit (Matrix B)

4) Pull of maximum joint profit against maximum difference (Matrix B)

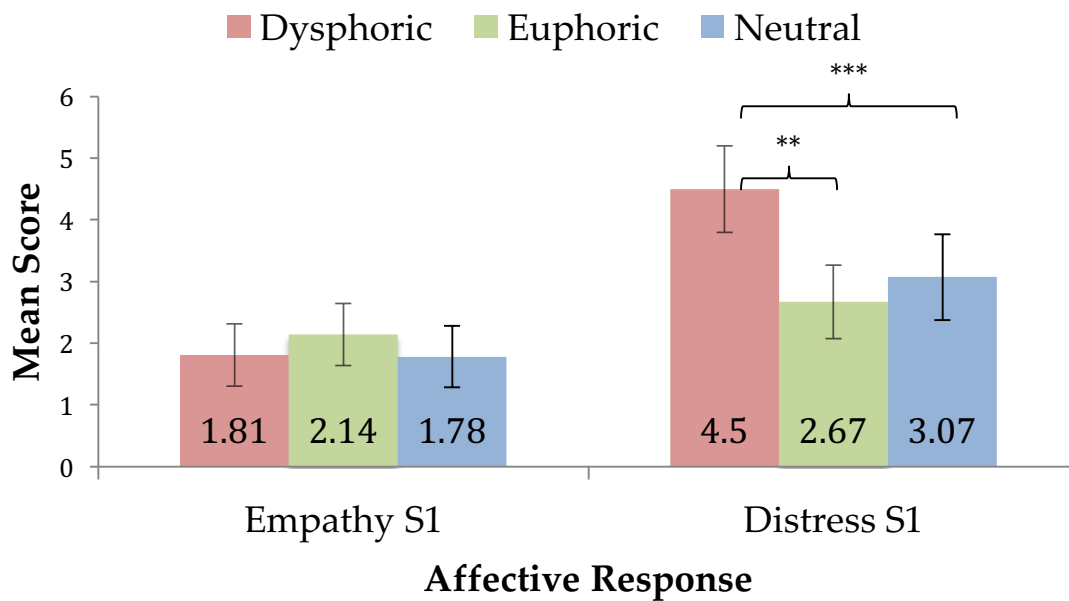
Manipulation Checks

To examine whether the experimental manipulations were successful in producing distinctive affective responses, both physiological and self-report measures were assessed. For physiological response, two post-manipulation HR measures were examined in a series of independent ANCOVAs⁵⁸. In these models, the post-manipulation HR measures were independently entered as outcomes, experimental condition was included as a predictor, and relevant baseline HR was controlled for as a covariate. In all of the HR analyses reported, a total of 35 responses were excluded due to errors with the HR monitoring equipment. Controlling for the baseline HR measure, experimental condition was not found to be significantly predictive of any difference during the ritual performance in either *average* heart rate, $F(2,69)=.457, p=.635$, or *maximum* heart rate, $F(2,69)=.251, p=.779$.

In summary, the physiological heart rate evidence was negative; the two measures collected revealed no significant difference between conditions. However, there was an unusually high amount of errors potentially because of the additional movement required in the ritual sequence and this may have contributed to unreliable results overall.

⁵⁸ There was an approximately linear relationship between pre- and post-HR measures for each experimental condition for both measures, as assessed by visual inspection of a scatterplot. Homogeneity of regression slopes was also found, as assessed by preliminary linear regressions that included interaction terms.

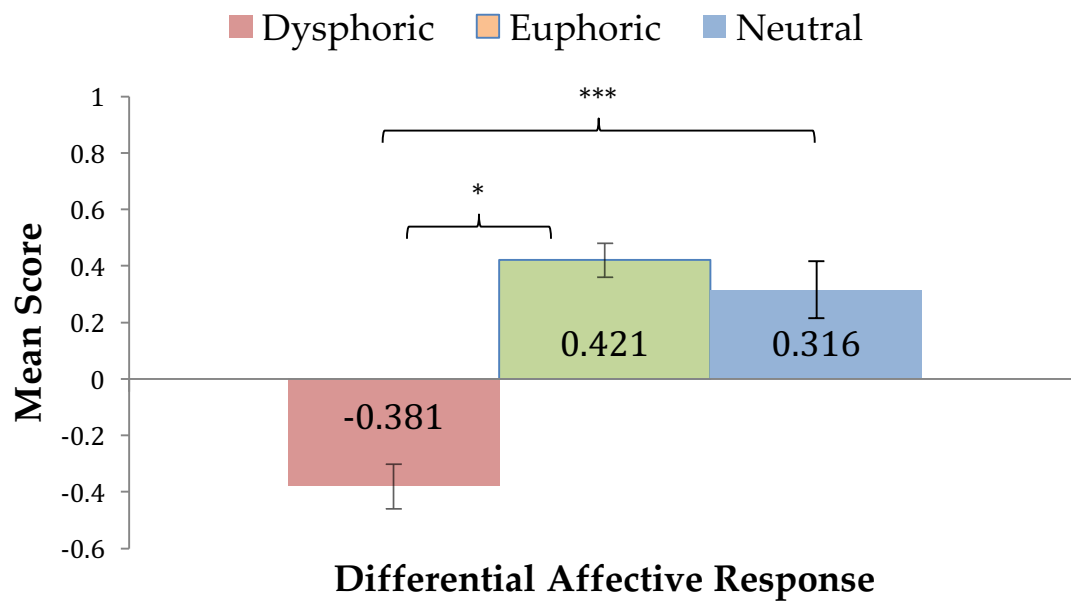
Figure 3-12: Affective response to infant sounds stimulus



Note: Error bars represent standard deviations.

A clearer difference between conditions was found when comparing Batson et al.'s (1983) self-report measures of empathy ($\alpha=.803$) and distress ($\alpha=.928$), in response to the baby audio stimulus. Partially replicating the findings from study 1, the crying baby in the dysphoric condition ($M=4.51$, $SD=1.75$) was found to elicit significantly stronger distress responses: $F(2,92)=13.873$, $p<.001$, and post-hoc analyses confirmed that this was significantly higher than both the euphoric ($M=2.667$, $SD=1.313$) ($p<.001$) and the control condition ($M=3.071$, $SD=1.585$) ($p<.001$) conditions. However, counter to study 1 there was no significant difference detected between conditions in empathetic response to the infant audio stimulus: $F(2,88)=1.268$, $p=.287$. The same patterns were replicated in session 2 which increased confidence in the validity of these findings.

Figure 3-13: Differential Affective response to infant sounds stimulus



Note: Error bars represent standard deviations.

Affective response between conditions were also compared using the aggregated positive ($\alpha=.740$) and negative ($\alpha=.794$) self-report responses on the I-PANAS-SF (Thompson, 2007). A pair of ANCOVAs were conducted on these measures to test for differences between conditions when controlling for baseline measures. There was no difference found for positive affect, $F(2,95)=1.489$, $p=.231$, but a significant difference was found for negative affect, $F(2,92)=3.111$, $p=.049$, $\eta^2=.067$. Subsequent pair-wise comparisons revealed that there were significant differences ($p=.049$) between the dysphoric ($M=3.150$, $SE=.133$) and euphoric conditions ($M=2.683$, $SE=.134$), but surprisingly no difference ($p=.432$) was detected between the euphoric and neutral conditions ($M=2.824$, $SE=.176$).

To further explore these findings, negative and positive affect scores were treated orthogonally and a differential affect score (positive-negative affect post-ritual) was calculated. Using this measure in an ANCOVA, controlling for differential affect in baseline

measures, significant differences were again detected between conditions, $F(2,91)=5.660$, $p=.005$, $\eta^2=.111$. Subsequent, pairwise comparisons indicated significant differences between the dysphoric condition ($M=-.381$, $SE=.177$) and both, the euphoric ($M=.421$, $SE=.177$) ($p=.006$) and neutral conditions ($M=.316$, $SE=.233$) ($p=.05$). However, again, with Bonferroni corrections applied, no significant difference was observed between the euphoric and neutral conditions ($p=1.000$).

Finally, comparable ANCOVAs were conducted with the independent affect and activation measures calculated from the single-item affect grid (J. A. Russell et al., 1989), but no significant differences in either affect, $F(2,93)=1.619$, $p=.204$, or activation, $F(2,93)=2.329$, $p=.103$, were found, when controlling for baseline measures.

In contrast to the HR data, and in accordance with the findings from study 1, the self-report measures provided strong support for the overall success of the manipulations. The dysphoric condition was clearly distinguishable, with significant increases in distressed responses and negative affect and, crucially, these responses demonstrated not only independence from the neutral condition, but also the euphoric condition.

It is less evident that the euphoric and neutral conditions were substantially distinct, as there were no significant differences in empathy or positive affect aroused. However, on almost all measures the distribution of the conditions accords with expectations with the euphoric condition displaying a high arousal, positive affect and high empathy profile and the neutral condition displaying a medium arousal, positive affect and medium to high empathy profile. Thus, the lack of significant differences may be a reflection of limited group sizes and a broader issue with how to effectively promote euphoria under the

current experimental conditions. Using a paired t-test on the PANAS baseline measures demonstrated that baseline levels of affect were significantly, $t(105)=7.72$, $p<.001$, more positive ($M=2.859$, $SD=.770$) than negative ($M=2.114$, $SD=.834$) which meant that it was easier to detect changes in the negative measure as there was greater scope for variation.

Hypothesis 1a (Dysphoric = More Fusion)

To assess whether the dysphoric ritual condition was associated with higher levels of identity fusion with the ritual ingroup a series of one-way ANCOVAs were conducted on the combined scores for the verbal fusion measure, while controlling for group identification.

Table 3-15: Descriptive statistics for group fusion and identification measures

Variable	Skew (SE)	Kurt. (SE)	α	M (SD)
Ritual Identity Fusion S1	.943 (.233)	-.128 (.461)	.939	2.425 (1.555)
Ritual Identity Fusion S2	.911 (.240)	-.375 (.476)	.901	2.399 (1.498)
Ritual Identification S1	1.037 (.241)	-.164 (.478)	.887	2.255 (1.496)
Ritual Identification S2	1.088 (.233)	.588 (.461)	.801	2.324 (1.463)

The suitability for an ANCOVA analysis was assessed through visual inspection of scatterplots, which confirmed there was a linear relationship between fusion and identification scores for each experimental condition, across both sessions. Homogeneity of regression slopes was also confirmed through non-significant interaction terms: S1- $F(2,94)=1.356$, $p=.263$; S2- $F(2,95)=1.219$, $p=.300$. Standardized residuals from each model were also normally distributed as assessed by Shapiro Wilks' tests ($p>.05$) and there was homoscedasticity and homogeneity of variances, as assessed by visual inspection of a scatterplot and Levene's test of homogeneity of variance (S1- $p=.457$; S2- $p=.718$),

respectively. Final examinations of residuals resulted in 2 extreme outliers (greater than 3 SDs) being excluded in both sessions.

The results from the ANCOVAs revealed no significant difference between conditions on identity fusion, when controlling for categorical group identification, in either session: S1- $F(2,93)=.077, p=.926$; S2- $F(2,92)=1.712, p=.186$. Post-hoc one-way ANOVAs, excluding group identification, also found no significant differences between conditions for either session. Paired t-tests contrasting fusion scores in session 1 with those from session 2 also demonstrated no significant difference in the scores between sessions: $t(97)=.666, p=.507$.

Hypothesis 1b (Dysphoric = More Relational Bonds)

Next, a pair of one way ANOVAs were conducted to test for differences between conditions on the combined measures of relational bonds in both sessions. Prior to analysis, inspection of boxplots indicated four cases as being extreme outliers. These were excluded as they were found to be exerting undue influence on the models across both sessions. With these cases excluded Shapiro-Wilks tests indicated normal distributions for all conditions across both sessions, except for the dysphoric condition in session 2 ($p=.005$). But as visual examination of histograms indicated only minor deviations from normality no further exclusions were made.

Table 3-16: Descriptive statistics for relational bonds

Variable	Skew (SE)	Kurt. (SE)	α	M (SD)
Relational bonds S1	.589 (.255)	.191 (.506)	.796	2.519 (1.154)
Relational bonds S2	1.010 (.255)	.914 (.506)	.890	2.396 (1.133)

The results obtained mirrored those reported for identity fusion. That is there was no significant difference observed between conditions on the combined relational bonds measures: S1- $F(2,90)=.281, p=.756$; S2- $F(2,92)=1.730, p=.183$.

Table 3-17: Spearman's correlation matrix for identity fusion and relational bonds

Variable	Rel. Ties S1	Rel. Ties S2	Emp.	Distr.	Pos (PANAS)	Neg (PANAS)	Age
Fusion S1	.533***	.579***	.334***	.141	.183	-.042	-.212*
Fusion S2	.566***	.800***	.344***	.128	.132	.017	-.216*

*= $p<.05$, **= $p<.01$, *** = $p<.001$

Post-hoc exploratory correlation analyses were conducted between ingroup identity fusion measures and similar variables to those explored in study 1, specifically: the relational bonds measures, age, measures of empathy and distress, physiological HR measures, and positive and negative affect scores. Using spearman's rho to correct for non-parametric distributions, no significant correlations were found with either HR measures. However, in accordance with expectations, significant positive correlations were observed between fusion and both relational bonds measures. Additionally, empathic response scores displayed a robust correlation with fusion measures, at both time periods (Table 3.18). These relationships were explored further in H3.

Hypothesis 2 (Dysphoric = Ingroup Preference)

Ingroup preference was assessed through three outcomes: 1) behaviour in an intergroup allocation task, using Tajfel matrices; 2) preferred choice of allocation target; and 3) implicit positive associations with group identity, calculated using positive adjectives and fictional ingroup/outgroup members in a BIAT.

Tajfel Allocation Task

As part of the experiment participants completed a series of decisions regarding the allocation of bonus payments to various targets including other members of their experimental ingroup and outgroup. The majority of these allocations were not intended for analysis but rather to disguise the purpose of the study and avoid demand characteristics. Analysis of post-experiment response sheets suggested that this had proven successful as, when asked their thoughts about the purpose of the experiments, although many participants speculated about the purpose of the ritual, none of the participants posited a link with the allocation tasks.

In all of the allocation tasks reported, the maximum that could be allocated to any individual over the four rounds was 820 円 and the minimum was 100 円. Normal distributions were found for all allocation outcomes and pull scores hence parametric comparisons were applied. All ANOVAs were checked for necessary assumptions and non-parametric comparisons were conducted to check marginal cases, however no significant difference in outcomes were found. Below then, the results from the original one-way ANOVA comparisons are reported.

Examining first the allocation choices made between a ritual ingroup member and outgroup member in session 1, there was no significant difference found, across the full sample, in total allocations to the ingroup member ($M=628.056$, $SD=267.873$) over the outgroup member ($M=601.111$, $SD=156.135$), $t(107)=1.030$, $p=.305$. Similarly, a one-way ANOVA on differences in allocations (ritual ingroup minus ritual outgroup donations) found no significant difference between conditions: $F(2,105)=.693$, $p=.502$.

A series of follow-up one-way ANOVAs comparing motivational pull scores also found no significant difference between conditions for the pull of: i) ingroup favouritism (FAV) over maximum joint profit (MJP), $F(2,105)=.394$, $p=.675$; ii) MJP over favouritism, $F(2,105)=1.203$, $p=.304$; iii) maximum difference (MD) over MJP, $F(2,105)=1.656$, $p=.196$; and iv) MJP over MD, $F(2,105)=.571$, $p=.567$.

Following these null results, the allocation choices made for the same targets in session 2 were examined. In contrast with session 1, this time a significant difference in total allocations made to the ritual ingroup member ($M=655.347$, $SD=152.089$ vs. ritual outgroup- $M=573.267$, $SD=130.001$) were detected when examining the full sample, $t(101)=5.223$, $p<.001$. However, when comparing conditions, the patterns replicated those found in session 1. That is, there was no significant difference between conditions in ritual ingroup – ritual outgroup donations, $F(2,98)=.320$, $p=.727$ and for the motivational pull of: i) ingroup favouritism (FAV) over maximum joint profit (MJP), $F(2,98)=.321$, $p=.726$; ii) MJP over favouritism, $F(2,98)=1.559$, $p=.216$; iii) maximum difference (MD) over MJP, $F(2,98)=1.065$, $p=.349$; and iv) MJP over MD, $F(2,98)=1.812$, $p=.169$.

In summary, allocations made on Tajfel matrices revealed no significant difference between experimental conditions on either the total amount of ingroup preferential donations or on the various intergroup motivations that Tajfel matrices can identify. This result contrasts from the marginal increase in cooperation observed in Study 1. Possible explanations for the discrepancy are explored in the discussion section.

Preferred Target of Bonus

At the end of the allocation tasks participants in S2 were given the opportunity to rank in order of preference the target that they would like the bonus payment based on their allocation choices to be sent if they could choose. This provided an additional opportunity to assess ingroup preference.

Across the full sample, a significantly higher rank was selected for ritual ingroup members ($M=3.18$, $SD=2.50$) over ritual outgroup members ($M=5.61$, $SD=2.178$) from the same experimental session, $t(100)=7.496$, $p<.001$. Additionally, the same pattern, with lower comparative ranks, was observed when comparing ritual ingroup ($M=5.37$, $SD=2.159$) and outgroup members ($M=7.58$, $SD=1.92$) from another experimental session, $t(100)=8.621$, $p<.001$. Comparing responses across conditions, a series of one-way ANOVAs were conducted on eight possible targets. However, no significant differences were found for rankings of i) a ritual ingroup member from another experiment, $F(2,97)=1.893$, $p=.671$; ii) a ritual outgroup member from the current experiment, $F(2,97)=1.214$, $p=.301$; iii) a ritual outgroup member from another experiment, $F(2,97)=.613$, $p=.544$; iv) a student from another university, $F(2,97)=1.965$, $p=.146$; v) an unknown Japanese participant, $F(2,97)=1.610$, $p=.205$; and vi) an unknown foreign participant, $F(2,97)=1.100$, $p=.337$.

The only targets that displayed a significant difference across conditions were the ritual ingroup member from the same experimental session, $F(2,97)=4.151$, $p=.019$, and a student from the same university, $F(2,97)=2.925$, $p=.058$. In the case of the ritual ingroup member, post-hoc bonferroni comparisons identified the observed difference to be driven by the higher rankings of the euphoric condition ($M=4.057$, $SD=2.899$) as compared to the

dysphoric condition ($M=2.421$, $SD=1.686$). There were no significant differences between conditions found in post-hoc analysis of the university student rankings.

In summary, in a ranking task, members of the euphoric condition were found to express a significantly greater preference for their fellow ritual ingroup members. This preference was not found to extend to any other category, including other categorical group members from other experimental sessions, increasing confidence in the validity of the result.

Implicit Positive Association

The results from the BIAT were analysed to examine implicit positive associations with assigned tribal group. Calculated D scores, with positive values indicating greater positive association with the ritual ingroup vs. the ritual outgroup, were found to be normally distributed, as assessed by visual examination of histograms and K-S tests: BIAT S1- $D(108)=.071$, $p=.200$; BIAT S2- $D(108)=.064$, $p=.200$. However, due to environmental differences in conditions between experimental rooms, independent t-tests were conducted with room as the between groups factor and a significant difference was found for both sessions: S1- $t(106)=-2.895$, $p=.005$; S2- $t(106)=4.537$, $p<.001$.

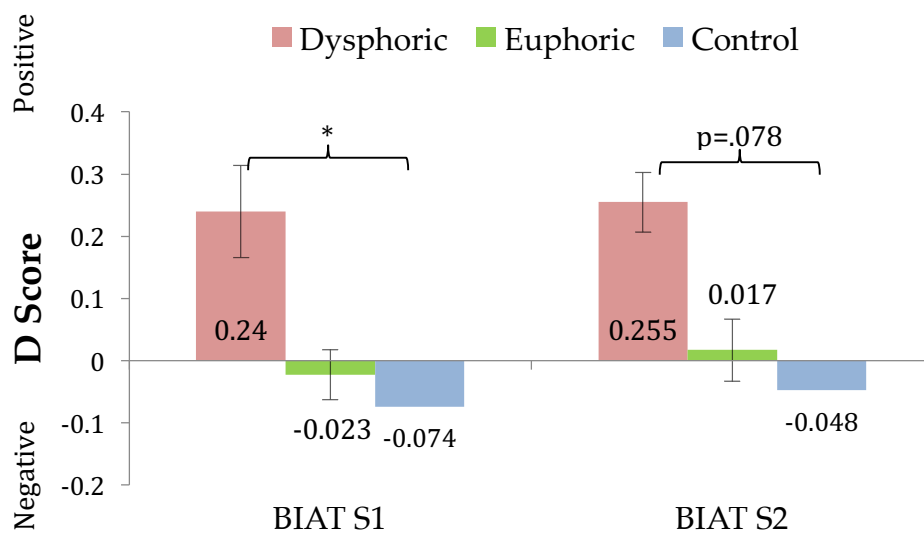
Table 3-18: Descriptive statistics for IAT outcomes

Variable	Skew (SE)	Kurt. (SE)	M (SD)
BIAT S1 (Positive Association)	-.124 (.246)	-.596 (.488)	.067 (.057)
BIAT S2 (Positive Association)	.100 (.246)	-.569 (.488)	.098 (.064)

Consequently, a two-way Factorial ANOVA was conducted for each session, with room (2) and condition (3) entered as separate fixed factors. Residual analysis was performed to test the assumptions of the two-way ANOVA and normality was assessed using Shapiro-

Wilk's normality test for each cell of the design. Homogeneity of variances as assessed by Levene's test was not violated in the models for either session (S1- $p=.851$; S2- $p=.794$). There were three outliers detected (one in S1; two in S2) as assessed by box-plot analysis. Two of these were found upon examination of session records to accord with potential interruptions from external noise. For consistency all three were excluded from analysis of the relevant session⁵⁹. Residuals for both sessions were found to be normally distributed ($p>.05$).

Figure 3-14: Implicit positive association with tribal names in both sessions



Note: Error bars represent standard deviations.

For BIAT responses in session 1, no significant interaction between condition and room was found, $F(2,101)=1.874$, $p=.159$, but there were independent main effects found for both room, $F(1,101)=11.585$, $p=.001$, $\eta^2=.103$, and condition, $F(2,101)=3.646$, $p=.030$, $\eta^2=.067$. Subsequent pairwise comparisons revealed that it was the difference between responses in the dysphoric condition ($M=.240$, $SE=.085$) and neutral condition ($M=-.074$, $SE=.097$) that

⁵⁹ For transparency purposes: All three outliers were in the dysphoric condition, but each was in a different session. Two cases were significantly lower than the mean and one was significantly higher.

was driving the effect ($p=.05$). There was also a difference detected between the dysphoric and euphoric ($M=-.023$, $SE=.087$) conditions but this did not reach statistical significance ($p=.099$).

For BIAT responses in session 2, again no significant interaction between condition and room was found, $F(2,100)=1.852$, $p=.162$, but there was an independent main effect found for room, $F(1,100)=7.589$, $p<.001$, $\eta^2=.201$, and a marginally significant main effect for condition, $F(2,100)=2.984$, $p=.055$, $\eta^2=.056$. Subsequent pairwise comparisons again suggested that the difference between responses in the dysphoric condition ($M=.255$, $SE=.089$) and neutral condition ($M=-.048$, $SE=.100$) was driving the effect but the difference this time was only close to statistical significance ($p=.078$). There was also a visible difference detected between the dysphoric and euphoric ($M=.017$, $SE=.090$) conditions but this again did not reach statistical significance ($p=.188$).

In summary, the BIAT results for both sessions indicated that respondents in the dysphoric condition had a significantly stronger positive association with their in-group tribal identity and fictional tribal members than those in the neutral condition. Results from the euphoric condition did not differ significantly from the neutral group but visual trends for both sessions suggested that there was a general stronger positive association. In session 1, immediately following the performance of the ritual sequence there was a marginally significant difference in positive associations between the euphoric and dysphoric conditions, but this was not replicated in session 2.

Hypothesis 3a (Positive Affect = More Fusion)

Based on the results from the BJJ Study and Study 1, it was anticipated that subjective positive assessment of the ritual experience, regardless of condition, would be correlated with higher levels of identity fusion. To examine this, simple pearson correlations were conducted between the fusion measures from S1 and S2 and the positive and negative affect measures. However, no significant correlation was found (all $p > .250$) with fusion in either session for either the positive or negative measure.

Table 3-19: Descriptive statistics for PANAS measures (post-ritual)

Variable	Skew (SE)	Kurt. (SE)	α	M (SD)
Positive (PANAS) S1	.152 (.244)	-.176 (.483)	.886	2.965 (.831)
Negative (PANAS) S1	.178 (.244)	-.071 (.483)	.851	2.894 (.966)

Hypothesis 3b (Positive affect = More Relational Bonds)

Subjective positive assessment was also anticipated to be associated with stronger relational bonds between participants. But simple pearson correlation analyses did not support this expectation as there were no significant correlations between relational bonds and the positive affect measures collected for either session. Negative affect in S2 displayed a negative correlation with relational bonds in S1.

Table 3-20: Pearson's correlations for relational bonds

Variable	Pos S1 (PANAS)	Neg S1 (PANAS)	Pos S2 (PANAS)	Neg S2 (PANAS)	Empathy	Distress
Rel. Ties S1	.184	.149	.092	-.237*	.358***	.032
Rel. Ties S2	.154	.117	.060	-.185	.385***	.128

*= $p < .05$, **= $p < .01$, ***= $p < .001$

In contrast with study 1, robust positive correlations were observed with the relational bonds and empathy measures but not with distress. To investigate these relationships and

those identified in H1 further, two hierarchical linear regression analyses were conducted. The first model included the relational bonds measure from session one as the outcome and the second model used the relational bonds measure from the second session. In both models the same hierarchical structure was employed: in the first stage, the empathy measure from session 1 was entered along with age as a control factor⁶⁰, and in the second session the identity fusion and group identification measures from session 1 were added.

Table 3-21: Hierarchical multiple regression on relational bonds S1

Variable	Outcome: Relational bonds S1			
	STAGE 1		STAGE 2	
	B (SE)	β	B (SE)	β
Constant	6.093 (1.548)	-	4.451 (1.372)	-
Empathy S1	.319 (.103)	.310**	.062 (.103)	.061
Age	-.243 (.072)	-.319***	-.179 (.064)	-.235**
Positive (PANAS)	.053 (.014)	.041	.051 (.113)	.115
Negative (PANAS)	.148 (.106)	1.397	.019 (.009)	.238
Identity Fusion S1	-	-	.165 (.081)	.238*
Group Ident. S1	-	-	.239 (.091)	.330**
R^2	.254		.456	
F	7.162***		11.468***	

$N=88$, * = $p < .05$, ** = $p < .01$, *** = $p < .001$, shading added to emphasis variables with significant coefficients.

Testing the variables and the two models generated for assumptions: first, examination of standardised residuals identified no cases as extreme outliers and hence no exclusions were made. Further examination of leverage and influence using Cook's and Mahalanobis distance values also indicated no other problematic cases. Both models demonstrated independence of residuals, confirmed through Durbin-Watson statistic values approaching 2 and partial regression plots also indicated suitably linear relationships for all variables. There was no indication of multicollinearity as assessed through examination of

⁶⁰ It was found to have a negative correlation with relational bonds in session 1 ($r = -.364, p < .001$)

correlations and VIF and tolerance statistics. Finally, examination of the residuals through histograms, P-P plots, and Normal Q-Q plots indicated appropriately normal distributions.

With relational bonds in session 1 as the outcome, the full model (Table 3.22) was able to predict a substantial portion of the variance in relational bonds; $R^2=.456$, $F(6,82)=11.468$, $p<.001$. The addition of identity fusion and group identification resulted in a dramatic improvement over the initial model, with the increase in explanatory power almost doubling; $\Delta R^2=.202$, $F(2,82)=15.228$, $p<.001$. Notably, the contribution of empathy disappeared in the second stage suggesting it was fully mediated by identity fusion and group identification. Positive and negative affect measures, as anticipated from the correlation analysis, did not significantly contribute to the model in either stage.

Table 3-22: Hierarchical multiple regression on relational bonds S2

Variable	Outcome: Relational bonds S2			
	STAGE 1		STAGE 2	
	B (SE)	β	B (SE)	β
Constant	3.209 (1.724)	-	.721 (1.153)	-
Empathy S1	.402 (.115)	.368***	-.053 (.086)	-.049
Age	-.103 (.081)	-.127	-.005 (.054)	-.006
Positive (PANAS)	.019 (.145)	.014	.051 (.113)	.115
Negative (PANAS)	.134 (.118)	.113	.095 (.077)	.080
Identity Fusion S1	-	-	.093 (.068)	.126
Group Ident. S1	-	-	.558 (.076)	.728***
R^2	.179		.659	
F	4.590**		26.401***	

$N=88$, $*=p<.05$, $**=p<.01$, $***=p<.001$, shading added to emphasis variables with significant coefficients.

In the next model with relational bonds in session 2 as the outcome, the full model (Table 3.23) was able to predict a substantial portion of the variance in relational bonds; $R^2=.659$, $F(6,82)=26.401$, $p<.001$. Again, the addition of identity fusion and identification in

stage 2 resulted in a dramatic improvement in explanatory power; $\Delta R^2=.480$, $F(2,82)=57.642$, $p<.001$. In this model however none of the other variables contributed significantly after the inclusion of the group identification measure. This suggests that group identification was the strongest predictor from the variables modelled and had maintained its predictive power after a week delay. The patterns observed in models for both outcomes increases confidence in the validity of the findings but also suggests that in this experiment it was group identification as opposed to identity fusion that was driving the formation of relational bonds.

The results from these regressions collectively did not support the hypothesised connection between subjective positive assessment of ritual experiences and relational bonds. Moreover, while fusion contributed independently to predicting relational bonds in session 1, its contribution was eclipsed in session 2 by group identification.

Hypothesis 4 (Dysphoric ritual experiences = Greater reflexivity)

In both sessions, participants were invited to write an open ended response about what they considered the meaning of the ritual performance to have been. These responses were rated by two hypothesis-blind, independent coders, who were asked to rate responses for i) the amount of introspective reflection displayed, using the same instructions as in Study 1, and ii) how symbolic the responses were, particularly in reference to the use of metaphorical comparisons. Ratings were completed using 6-point scales with the higher values indicating greater reflexivity and more symbolic interpretations. Raters also independently identified 11 participants that had shown no engagement with the question

and focused exclusively on experimental features as opposed to the ritual. These 11 responses were included for reliability tests but excluded from other analyses.

Comparing results from the two raters, correlations were high for reflexivity and symbolic ratings across both samples (Table 3.25). Inter-rater agreement ratings (Landers, 2015), calculated using intraclass correlations (Shrout & Fleiss, 1979), also indicated an excellent level of agreement for reflexivity and symbolism (Table 3.25). The two ratings were thus averaged for each session into a single score for the two measures and difference scores were calculated for the values between S1 and S2.

Table 3-23: Reliability of SER Ratings

Variable	<i>r</i>	ICC (2,2) [95% CI]	α
Reflexivity S1	.889***	.939 [.909, .958]	.941
Symbolism S1	.831***	.892 [.813, .934]	.907
Reflexivity S2	.795***	.881 [.802, .925]	.895
Symbolism S2	.810***	.881 [.826, .919]	.886

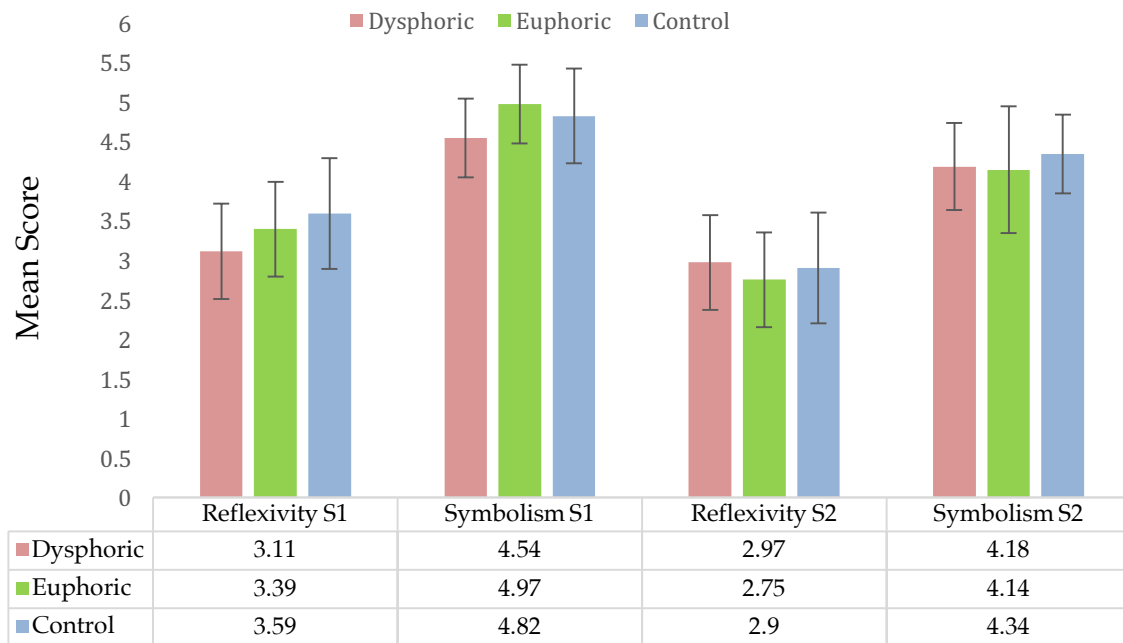
N=108, $p < .05^*$, $p < .01^{**}$, $p < .001^{***}$

Comparing responses by condition, first a series of exploratory correlation analyses were conducted to identify any demographic or contextual factors that could confound results. Age was the only significant demographic factor identified and it was found to have a weak but significant negative correlation with all measures except symbolism in S1. Experimental room was the only significant contextual factor with a weak correlation with reflexivity in session 2. Both factors were entered as controlled covariates in a series of ANCOVA analyses reported below.

Controlling for age and experimental room, experimental condition was not found to be significantly predictive of any difference in SER scores across both sessions: Session 1- *reflexivity*, $F(2,79)=.077$, $p=.926$, *symbolism*, $F(2,79)=.163$, $p=.850$; Session 2- *reflexivity*,

$F(2,87)=1.058$, $p=.352$, *symbolism*, $F(2,87)=.608$, $p=.547$. Examining changes in reflexivity, $F(2,79)=1.227$, $p=.299$, and *symbolism*, $F(2,79)=1.482$, $p=.233$, scores between sessions also revealed no difference between conditions.

Figure 3-15: Ratings of reflexive accounts of promotional experiences



$N=97$, $p<.05^*$, $p<.01^{**}$, $p<.001^{***}$

Again the results obtained are counter to the findings of Richert et al. (2005), and although a different rating scheme was applied there was an additional effort made through the inclusion of the symbolism score to approximate a measure of analogical depth. An actual precise comparative measure of analogical depth was not employed as the thesis author was unable to locate any suitable measure for Japanese language sources. Both the reflexivity and the symbolism scores demonstrated moderate declines between session (Fig.3.15).

Table 3-24: Spearman's correlations for reflexivity and physio./affect. measures

Variable	Max HR	Avg. HR	Pos. Aff S1	Neg. Aff S1	Fusion S1	Fusion S2	Group Ident. S1	Group Ident S2
Reflex.S1	-.180	.173	.053	.141	.169	.093	.080	.149
Reflex.S2	-.149	-.061	.217*	.219*	.242*	.179	.145	.245*
Symbol. S1	.016	-.192	.183	-.024	.126	.087	.130	.177
Symbol. S2	-.112	-.003	.135	.018	.146	.095	.117	.115

N=88, *= $p < .05$

Based on the findings of Richert et al. (2005) possible connections between physiological arousal, measured using collected HR metrics, and the *reflexivity* and *symbolism* scores were also examined but no significant correlations were found. Replicating the analyses conducted in Study 1, there were also no significant correlations detected between *reflexivity* scores in session 1 and identity fusion measures and group identification measures or with self-reported post-ritual measures of affect. For session 2 however both the positive and negative affect measures demonstrated a significant positive relationship with the *reflexivity* measure, as did fusion (from session 1) and group identification (from session 2). There were no significant correlations found with *symbolism* measures in either session.

Table 3-25: Multiple regression on reflexivity S2

Variable	Outcome: Reflexivity S2	
	STAGE 1	
	B (SE)	β
Constant	.802 (.645)	-
Positive (PANAS)	.201 (.170)	.122
Negative (PANAS)	.310 (.144)	.218*
Identity Fusion S1	.229 (.124)	.260*
Group Ident. S2	-.055 (.133)	-.058
R^2	.128	
F	3.143*	

The inconsistency of the results between sessions for the *reflexivity* score suggests that as previously hypothesised that the impacts of introspective reflection require time to develop. This accords with the findings of Richert et al. (2005) and the overall theoretical model in Modes concerning the importance of spontaneous exegetical reflection (SER) in generating deeper group bonds. Finally, an exploratory linear regression based on the observed correlations (Table 3.26) was conducted with reflexivity in session 2 as the outcome and all significant correlating variables entered simultaneously as predictors. The model that emerged demonstrated, in accordance with theoretical models, that post ritual negative affect and identity fusion proved to be the only significantly predictive factors of greater reflexivity.

3.7 Discussion

This study represented an attempt to refine and expand the collective artificial ritual methodology applied in Study 1, and in so doing provide an independent test of the previous results. The results obtained, as with study 1, are mixed but there are a number of areas of consistency that increase confidence in the validity of the previous findings.

One clear area of agreement, concerns the lack of evidence for a significant increase in fusion (H1a) or self-reported relational bonds (H1b) with other group members in the dysphoric condition. The null results obtained also extended to the euphoric condition and therefore suggest that this is not just a case of dysphoric experiences being so unpleasant in the short term that they obscure smaller shifts in perception. Furthermore, this study also made use of verbal scales for identity fusion and group identification, which are typically considered more valid than the visual measures used in study 1. However, rather than

calling into question the connection between dysphoric experiences and identity fusion, which has been attested to in a number of studies with greater ecological validity (Buhrmester et al. 2014; Atran et al. 2014b; Whitehouse et al. 2014; Whitehouse et al. *submitted*), the current results demonstrate that exploring a construct like identity fusion may not be possible through the bonds forged by temporary experimental groups. This is likely why the majority of successful identity fusion studies to date have made use of existing group identities in their manipulations.

Yet, the results from H2 also suggest that some caution is needed before research avenues are pre-emptively closed. Specifically, although the primary measure of ingroup preference displayed no difference between conditions, both of the subtler outcome measures- the BIAT and the preferred target for allocation measure, which suffer less from potentially confounding demand characteristics, were significantly different between conditions. In particular, the BIAT indicated that there was a substantially greater positive association with hypothetical tribal members in the dysphoric condition, while the preferred target measure indicated a stronger preference for the other experimental ingroup members. Thus, although it remains unlikely that a novel experimental group lacking any substantial emotional connection to the other participants could engender a strong sense of identity fusion amongst members, some of the necessary precursors that lead to fusion may be identifiable.

To elaborate, the implicit association measure involved participants classifying fictional members of their tribal ingroup, alongside members of an equally fictional tribal outgroup using a simple naming heuristic. The speed of these classifications was then compared

when combined with a similar classification task involving positive and negative adjectives. The theory being that classifications will be quicker and easier when congruent categories are paired and slower and more difficult with incongruent categories. For the current experiment if a preferential association existed for the tribal ingroup then pairings of ingroup tribal names and positive adjectives should be completed quicker and more accurately than for outgroup tribal names. This pattern of ingroup preference was detected for participants in the dysphoric condition across both sessions and was of a substantially greater magnitude than observed for the control and euphoric condition.

On the face of it, this is a counterintuitive finding, as it indicates that individuals, whose only exposure to a particular group identity was through performing a stressful task while listening to a screaming infant, hold positive associations with that identity. However, this partly accords with the predictions of Modes theory on the basis that shared dysphoric experiences are expected to produce salient memories and encourage ritual performers to feel a stronger kinship with those who have similar experiences. One problem for this interpretation is that the tribal members used for the BIAT were fictional and did not include the names allocated to the other members of the experimental ingroup, who had just performed the same task. There is thus some ambiguity over whether it is reasonable to infer that the preference displayed relates to an expectation of shared experiences with fictional tribal members. A speculative explanation would be that this represents preliminary evidence for the projection processes needed to establish the bonds posited by *extended fusion*, whereby relational bonds are projected onto a categorical group identity (Swann et al., 2012). Ultimately, the present study is unable to establish whether the result

of the BIAT is based on a positive association with a categorical group identity the tribal group or a positive association with fictional group members which would represent a relational target, so this is an issue that deserves attention in future research.

Cognitive dissonance (Festinger, 1962) could be invoked as an alternative explanation but first, most research on dissonance relates to efforts required to join a group (Aronson & Mills, 1959; Gerard & Mathewson, 1966; Hautaluoma & Spungin, 1974; Lodewijkx & Syroit, 1997) rather than a task performed as a member of a group. Second, dissonance accounts tend to present dissonance as being a semi-conscious process, in that it involves cognitive assessments of group commitments. This is an issue when it comes to accounting for the observed results as the core design of an implicit association task is that it is intended to bypass reflective cognitive processes and identify *implicit* cognitive associations. This is one of the strengths of the IAT measure and the fact that it found a clear difference between conditions while the primary allocation task did not, emphasises the importance of considering implicit processes when designing experimental protocols.

The second significant measure of ingroup preference was the ranked allocation preference item. In contrast to the BIAT results elevated preference for ingroup targets was found in the euphoric condition. Additionally, it was only relational ingroup members from the present experiment that were significantly more highly ranked which suggests that no significant preference was extended to categorical ingroup members from other experiments. The lack of preference displayed towards categorical group members is also congruent with the lack of implicit positive association for fictional group members in the

BIAT. However, this assumes that participants did not project relational ties onto the fictional members, which was proposed as a possible interpretation of the results above.

Against this interpretation is the lack of any stronger preference for relational ingroup members from their experiment group displayed by those in the dysphoric condition. This adds support to the view that the dysphoric condition did promote affiliative tendencies that were more categorically orientated, which is counter to theoretical predictions.

In summary, these mixed results complicate any straightforward interpretation and also raise concerns about the reliability of the results. If the results are valid then the inference is that both the high arousal conditions lead to heightened affiliative tendencies but the euphoric condition promoted relationally focused ingroup preference, whereas the dysphoric condition encouraged positive association with categorical members. This is the converse of the connections proposed in Whitehouse & Lanman (2014) concerning euphoric and dysphoric arousal and the type of group bonds they forge. Yet, the results do accord with broader theoretical expectations that euphoric and dysphoric arousal promote group bonding in distinctive ways (Kelly et al., 2014; Y. I. Russell et al., 2011).

Returning to the null result with the Tajfel tasks, an important distinction between the tasks used in Study 2 and Study 1 is that Tajfel matrices are primarily focused on intergroup interactions, whereas the PGG in Study 1 was focused on cooperation within an ingroup. It is possible therefore that the null finding is due to the fact that shared dysphoric or high arousal ritual experiences more generally elevate ingroup cooperation but do not promote intergroup discrimination. There is an ongoing debate in psychology about the extent to which ingroup preference necessarily entails outgroup discrimination and

identity fusion theory suggests that it is salient group norms that determine the character of pro-group behaviour (Buhrmester et al., 2012; Semnani-Azad et al., 2012). Overall, the tests for H2 provided mixed results that require further investigation.

A clearer signal was found for the predictive power of the novel *group identification* measure. Focusing on the *relational bonds* measure from session 1, fusion and identification were found to both act as significant, independent predictive factors. But for the same relational bonds measure one week later in session 2, group identification proved to be the sole predictive factor in a model accounting for a full 65.9% of the variation in the measure. This was a particularly unexpected finding given the clear association between fusion and relational bonds in the literature and in the BJJ data.

The identification measure was based on two well validated identification scales and, aside from the contradictions the finding poses with the overall identity fusion literature, there seems no reason to doubt its veracity. This means that it may be that the bonds created through the experimental manipulation were of a more categorical than relational nature. Yet, there was no significant difference in identification measures detected between conditions in either session: S1- $F(2, 105)=.803, p=.451$; S2- $F(2,98)=1.733, p=.182$, so even if this were correct it cannot alone account the outcomes observed in H2. What the finding does illustrate clearly is how important it is to include control measures for alternative categorical pathways to affiliation if the intention is to to make clear statements about a specific role for fusion-like bonding. The lack of any group identification measures in the streamlined methodology of Study 1 thus represents a significant limitation and calls into question the broader applicability of the previous findings.

Finally, in relation to H4 and predictions about the increased reflexivity engendered by dysphoric experiences and elevated physiological arousal, only limited supporting evidence was provided. Overall, the results replicated the null findings from Study 1, despite including an additional measure of *symbolism* that was intended to better reflect the ‘analogical depth’ measure employed by Richert et al. (2005). However, in post-hoc regression analyses of the delayed SER measure which is argued to be a more ecologically valid test of the Modes predictions, negative affect was found to be a predictive factor, along with the ingroup identity fusion measure but the proportion of variance in reflection accounted for by these variables remained small ($R^2=.128$).

Conclusions

The two studies reported in this chapter served to illustrate the potential for employing an artificial ritual methodology to examine specific features of ritual performances. In particular, groups of strangers were found to increase their cooperative tendencies (Study 1) and to display more positive associations with fictional ingroup members (Study 2) after enduring a collective dysphoric ritual. Crucially, the inclusion of a euphoric condition in Study 2 lends credence to the claim that such effects are not being driven by higher arousal or affective stimulation more broadly. However, in a preference ranking task an elevated preference for other relational ingroup members was observed only in the euphoric condition. This finding complicates any straightforward interpretation of the results and raises doubts over whether the results observed with the dysphoric condition in Study 1 would be restricted to shared negative affective experiences.

Overall, it is important not to overstate the solidity of the findings from these studies as while pretesting in Study 2 did establish affective manipulations that produced responses that could be statistically distinguished into three categories (neutral, euphoric, and dysphoric). Disentangling the neutral and euphoric condition proved difficult in manipulation checks. Future efforts need to be made then to further promote more distinctive affective manipulations. However, it is also possible that distinctions between neutral and positive affect in an artificial ritual setting will pose very difficult to establish particularly since baseline affect is rarely neutral.

The combined results from both studies also failed to find the predicted associations between the dysphoric ritual conditions and both elevated levels of fusion and perceived relational bonds. Given the consistent negative findings across a two country sample, this suggests that a ritual performed in a laboratory setting with a group of strangers may not be an appropriate means for generating the intense bonds characteristic of identity fusion. Future studies might instead focus on exploiting existing group memberships or personally consequential memories as an alternative, more fusion relevant manipulation. This is not to argue that the artificial ritual paradigm should be abandoned as this would be premature on the basis of the current studies and, moreover, null results should not necessarily be taken as the result of a failure of a methodology.

The significant results obtained from the BIAT and allocation ranking preference task also underscore the need for future experimental ritual methodologies, particularly those that involve naïve participants and novel groups to explore subtler, unconscious cues of

increased group affiliation. Self-report items should not be excluded but instead used in combination so that they can also serve as independent validity tests of key findings.

Limitations

Both of these studies relied almost exclusively on student participants and hence the generalisability of the findings to broader populations are limited. Furthermore, Study 2 employed a single sex sample. This removed potentially confounding interactions based on gender and is also reflective of real rituals, which are often gender restricted, but it reduces the relevance of the reported findings to potentially 50% of the population.

The significant methodological changes introduced between studies and the cross country samples also limit how far comparisons between findings can be considered valid. In particular, Study 1 employed an ingroup cooperation measure as its key behavioural outcome, whereas Study 2 used an intergroup discrimination task, which could account for the variance in findings. Ingroup preference is not always accompanied by intergroup discrimination and thus it may be that an ingroup cooperation measure would have produced more consistent results.

It is also possible that the differences observed in outcomes are due to cross-cultural factors but the methodological differences between studies mean that it is impossible to ascertain this on the basis of the current evidence. In the future, it would thus be desirable to replicate the methodology from Study 2 with a comparable UK sample to examine whether the findings from Study 1 were replicated and whether cross-cultural factors contributed to different outcomes in Study 2. In order to facilitate such comparisons, it

would be crucial to take additional personality measures, such as idiocentric and allocentric tendencies, to help distinguish cultural and individual level differences.

Furthermore, although the methodological changes introduced in Study 2 did successfully disguise the relevance of the behavioural allocation tasks they also significantly complicated the design of the study. Future studies attempting to make use of the core methodology may be better served to adopt a design that falls somewhere between the two protocols presented above.

Finally, physiological measures in both sessions proved problematic with frequent loss of signal and potential issues with the validity with the data collected. More reliable means of collecting such measures are available but they all carry the risk of becoming the most salient feature of the experiment and care must be taken to balance the need for greater precision with the level of intrusion required. In both studies, even with various efforts made to reduce the salience of the physiological measures, a substantial proportion of participants focused on heart rate and physiological arousal when discussing the significance of the ritual.

CHAPTER 4: DYSPHORIC FIREWALKING RITUALS

Chapter four is based on questionnaire data collected at two extreme fire walking festivals performed at Shinto Shrines in small fishing communities located in Northern Japan. The questionnaires were designed to serve as a conceptual replication of recently reported findings from dysphoric rituals in Mauritius (Xygalatas et al., 2013b). Structurally this chapter is organised around a quasi-experimental field study that relied on pre- and post-event data collection periods to explore the impact of observing or performing in dysphoric fire walking rituals. The chapter begins with the theoretical rationale for the study and provides some background detail on why the specific sites in Shakotan and Furubira were selected. The demographic and social context of the field sites are briefly described, before the methodology of the questionnaire and the results collected from performers and observers are reported and analysed. Primarily, this chapter addresses whether ritual participation, or observation, is associated with increases in identity fusion with a variety of targets, the endorsement of more exclusive personal identities, and increases in prosocial attitudes and behaviour. Self-report and behavioural measures of willingness to engage in costly sacrifice in keeping with the studies reported in the previous chapters were also employed.

4.1 Rituals in Japan

When compared to Western norms, Japanese society is often presented as highly ritualistic and group orientated. Such portrayals are found not only in popular books (Garcia, 2011; Norbury, 2011; Zimmerman, 1985) but also in many ethnographic studies (Ben-Ari, Moeran, & Valentine, 1990; Benedict, 2006; Dore, 1973; Moeran, 1984; Nakane,

1973). Similarly, in the cross-cultural psychology literature Japan has often been featured as the paradigmatic collectivist society contrasted against American individualism (Oyserman et al., 2002; Yuki & Takemura, 2014; Yuki, 2003).

There have been many critiques, particularly from anthropologists, of group orientated models of Japanese society (Befu, 1980, 2001; Ben-Ari et al., 1990) and associated portrayals of Japan as a singular national stereotype of homogeneity, conformity, and hierarchy (Befu, 1992, 2001; Ben-Ari et al., 1990; Vlastos, 1998). However, there is less criticism directed at the significance afforded to ritual in Japanese cultural and social life. In the anthropological literature Japan is recognised as possessing an elaborate and complex ritualized gift exchange culture (S. E. Beatty et al., 1996; Befu, 1968; Minowa, Khomenko, & Belk, 2010; Rupp, 2003) and the associated aesthetics of gift-wrapping have even been proposed as a suitable metaphor for understanding the multi-layered nature of modern Japanese society (Ben-Ari et al., 1990). Furthermore, across a variety of group contexts there is a recurring tendency for collective dysphoric events to feature as an important ritualised aspect of group activities, particularly in relation to initiation events and religious festivals. The association is not restricted to religious groups (Anderson Sawada, 1998; Lobetti, 2013; Miyake, 1989; Reader, 1996; Stevens, 1988) but also includes other social groups such as university clubs and dormitories, martial arts groups (Chapman, 2005; Pettas, 2012; Twigger, 1999), and the hazing-like drinking activities associated with sports societies and commercial activities in Japan (Allison, 1994; McDonald & Sylvester, 2013; Moeran, 2005; Yoshihara & Shimizu, 2005).

In terms of formal ritual events, Japan is characterised by the presence of a vast array of national and local ritual festivals (*matsuri*) organised largely independently and conducted primarily by Shinto shrines and Buddhist temples (Covell, 2005; Nelson, 1996a, 1996b; Reader & Tanabe, 1998; Reader, 1990). Buddhism, Shintoism, and to a lesser extent, Christianity, are the major doctrinal religions present but there are also a wide variety of local folk beliefs and practices which are often incorporated in festival celebrations. Non-exclusive religious affiliation is also a cultural norm in Japan with the different religious systems serving complementary functional roles, as captured by the phrase 'Born Shinto, Marry Christian, Die Buddhist' (Reader, 1990: 7) which reflects the common functional "division of labour" between the religions (Mullins, 2011: 67).

Prior to reforms in the Meiji period, Buddhism and Shinto were also less strongly distinguished and there was a popular religious philosophy (*honji suijaku*) that presented the Shinto deities (*Kami*) as being local manifestations of Buddhist deities. Combined Temple-Shrine structures were also common and despite government mandated separation during the nationalist Meiji era some examples of this overt synthesis continue to persist into the modern era. Christianity, while remaining a minority religion in Japan with less than 3% of the total population (NHK Culture Research Institute, 2008) is popular for marriage ceremonies (Mullins, 2011; Reader, 1990). There are also a variety of New Religious Movements (NRM) in Japan, many of which rose to prominence during a religious boom referred to as the 'rush hour of the Gods' in the 1970s (Hardacre, 1988; McFarland, 1967; Shimazono, 2004). These groups remain a significant feature of the Japanese religious landscape and have also influenced mainstream culture as demonstrated

by the Komeito an influential political party which was founded by members of the Buddhist NRM Soka Gakkai. However, the popularity of NRMs declined significantly following the Tokyo subway attack, wherein poisonous sarin gas was released by followers of a millennial NRM called *Aum Shinrikyo* in 1995 (Hardacre, 2007; Reader, 1996).

Aside from practices performed at household altars (Nelson, 2008), the majority of Japanese do not perform regular private religious ritual practices (Mullins, 2011; Reader, 1990). A point which is reflected by Japan frequently being ranked as one of the least religious countries in the world (WIN-Gallup International, 2012; WinGallup, 2014). Japan's position in these polls of religion is partly due to culturally specific connotations of terms like religion, which for Japanese people mean that when asked questions about 'religion', 'faith', or 'God' they tend to associate the terms with the concept of Christianity: a religion that has specific doctrines and requirements for church membership, or with devout religious adherents. By these standards only around 33% tend to report having personal beliefs however when questions are phrased in relation to less formal terms like 'spiritual beliefs' or participating in religious actions such as festivals, annual events and rites of passage the results are very different (Mullins, 2011: 65). In fact, counter to often cited official statistics using these more flexible metrics Japan could actually be categorised as a country of "mass religiosity" (Mullins, 2011) as almost all Japanese people attend shrines and temples at several times during the year and participate in a variety of festivals.

Nelson (1996b) describes participation in such events as being "a very modern practice" defined by "freedom of expression" in regards to motivations and interpretation of meanings. In the typology of Modes theory almost all of these events would be categorised

as imagistic rituals because they are typically performed once a year, involve little or no official dissemination of systemised doctrinal teachings, and tend to feature various high-arousal performances; from intense drumming and dancing (*Yosakoi*) to firewalking (*Hiwatari*) (Fig.4.1) and immersion in ice cold water (*Misogi*) (Fig.4.2).

Figure 4-1: A Buddhist firewalking festival held in Miyajima



Source: *Photo taken by author*

Reader & Tanabe (1998) argue that all of the religious systems in Japan combine to form a syncretic “common religion” of Japan that is focused on practical participation with religious systems in order to gain “this-worldly benefits” (*genze riyaku*). This claim has been supported by a number of ethnographies that detail the lives of priests and their families in Buddhist and Shinto institutions (Covell, 2005; Jaffe, 2001; Nelson, 2000) and it also echoed by the Japanese scholar Ama Toshimaro who wrote a popular book titled ‘*Why are the Japanese non-religious?*’ (*Nihonjin wa naze mushūkyō na no ka*) (Ama, 2005). Ama suggested

that a distinction needed to be made between “revealed” religions, which have founders, doctrines and institutional membership, and “natural” religion, which represents the non-denominational religion of the vast majority of religious people and argued that for most Japanese being “without religion” (*mushūkyō*) simply refers to not believing in one single religion rather than being non-religious.

Figure 4-2: A Shinto ice-water immersion festival held in Tōkyō



Source: Image used with permission of Yoshio Wada

In addition to the popular religious activities participated in by the majority of the Japanese population, a smaller minority self-identify as religious and engage in regular devotional and self-cultivation practices. These practitioners are typically attached to a local shrine or temple and often belong to lay associations of devotees and volunteers. Such

groups are instrumental in the organisation of large ritual events (Nelson, 1996a) and also provide support to individual ascetics (Ludvik, 2006). Individual ascetic practices are performed by a small minority of individuals but there are a variety of ascetic traditions that exist both within and outside of institutional religious groups (Lobetti, 2013; Miyake, 1989) and despite the low number of practitioners these traditions have established numerous well-known cultural motifs, such as purifying oneself by standing under rushing waterfalls.

From the above summary it should be apparent that despite its reputation as an atheist nation Japan provides a rich environment in which to explore ritual activity. Additionally, due to the practice centred cultural context it also provides a unique opportunity to address the impact of rituals and ritualised behaviour when overt devotion to a particular religious doctrine or identity is largely absent.

4.2 Rationale for Study

The studies reported in the previous chapters examined collective dysphoric ritual events in the context of martial arts promotion ceremonies (Chapter 2) and in artificial rituals conducted in a controlled experimental setting (Chapter 3). Both of these accord with the broad definition of ritual as outlined in Chapter 1, however neither is particularly comparable to a religious ritual context. The artificial ritual contained procedural elements that have symbolic parallels with religious symbolism (e.g. anointing the wooden idol) but the entire ritual performance was deliberately designed to avoid any specific connotations with any known culture or religion. This thesis has argued in favour of a broad understanding of ritual, and the benefit of fractionating specific aspects of ritual and

ritualised behaviour for analysis independent from religious contexts however it is also the case that religion is widely recognised as a paradigmatic domain of ritual practices (Purzycki et al., 2014; Purzycki & Sosis, 2013; Rappaport, 1999). Additionally, most CSR ritual theories, including Modes and ritual costly signalling theories, developed from research that was primarily focused on religious rituals (J. L. Barrett & Lawson, 2001; Bulbulia & Sosis, 2011; Whitehouse, 1996, 2004). As such it is useful to examine the consequences and processes underlying collective rituals performed in a genuine religious setting.

There are a variety of features that distinguish religious rituals from secular versions and chief amongst these are references to supernatural agents or forces and sacred ideologies (Boyer, 2001; Rappaport, 1999). Sosis & Bressler (2003: 229-230), for instance, argue that although all costly rituals including secular ones may “generate emotional experiences that engender a sense of community amongst performers”, religious rituals uniquely possess the ability to imbue ritual events with a sense of sanctity and connect them with unfalsifiable beliefs that make the bonds they establish less ephemeral. The literature on religious priming studies also suggests that reminding people of religious concepts can generate prosocial behaviour and attitudes, at least amongst believers (Batara, 2015; Norenzayan et al., 2013; Shariff & Norenzayan, 2007; Shariff et al., 2015). Similarly, a recent field experiment conducted in Mauritius found that participants in a bargaining game were more cooperative when participating in a religious (temple) setting as opposed to at a restaurant irrespective of stated religiosity (Xygalatas, 2013a). However, despite the studies and meta-analyses cited above it would be inaccurate to claim that there is a

complete consensus that religious rituals are uniquely effective at forging long lasting bonds, or that religious concepts and contexts are more broadly associated with prosociality. There remains debate in the field with failed registered large scale replication studies (Gomes & McCullough 2015) and core methodological criticisms made of individual papers (Randolph-Seng & Nielsen 2008) and broader meta-analyses (Van Elk et al. 2015). Furthermore, many of the studies that posit a unique role for religious rituals do so by arguing that rituals naturally complement strong religious beliefs, but as discussed this does not reflect the religious context in Japan where strong devotional beliefs are rare. Ultimately then the distinctive nature of religious rituals and whether they produce effects beyond rituals from non-religious contexts remains unresolved and in need of further investigation.

This chapter seeks to contribute to these ongoing efforts by testing the key predictions of CSR theories concerned with dysphoric rituals in a religious context and comparing the findings against those of previous chapters. However, it should be cautioned that due to the clear methodological and population differences any such comparisons are not intended to be definitive but rather to offer preliminary evidence that would need to be examined in further studies before any generalisations could be made with any confidence.

Despite these caveats, exploring collective dysphoric rituals in a real religious context increases the ecological validity of the results and enhances the potential for the ritual to be personally consequential for some of the performers and observers involved. The ritual performances examined in this chapter are all part of Japanese festival (*matsuri*) celebrations which are conducted at various Shinto shrines and Buddhist Temples. These

kinds of religious festival events are ubiquitous throughout Japan, but for the purposes of this thesis, less common extreme firewalking festivals (*Hiwatari matsuri*) were sought out. Two festivals were selected as suitable and these were both performed by local Shinto shrines in two small fishing villages (Furubira and Shakotan) located in Hokkaido the Northern island of Japan. These festivals were selected due to the firewalking festivals performed being of a more extreme and dysphoric nature than is typical for firewalking events, as they involve individual performers and groups carrying heavy wooden portable shrines (*mikoshi*) passing directly through burning flames as opposed to the more common practice of walking barefoot over burning ashes. The festivals also both occurred within two neighbouring towns and were performed over multiple nights which provided additional opportunity for collecting responses. The fact that the two festivals were Shinto events should not be interpreted as demonstrating that Shintoism in general is more closely associated with extreme rituals than Buddhism in Japan as there are a large number of comparable dysphoric Buddhist matsuri.

As discussed above, conducting the studies in Japan presents the unique opportunity to examine dysphoric collective rituals that are performed in the widespread absence, amongst performers and observers, of committed doctrinal beliefs. This does not make festivals in Japan non-religious events but rather highlights the deficiencies and potential ethnocentric assumptions inherent in definitions of religion that focus solely on institutional belief. Japanese festivals are organised by religious authorities and associated lay volunteers, conducted primarily in and around Shrine and Temple facilities, and contain a plethora of religious symbolism. That the people who attend them, do not

typically identify as religious members of any community, or endorse statements about monotheistic deities (WinGallup, 2014) is not surprising in a culture where syncretism is the norm (Mullins, 2011; Reader, 1990) and 'religion' holds associations with state mandated worship (Earhart, 2005; Hardacre, 1986; Harootunian, 1999; Victoria, 2006) and dangerous religious cults (Hardacre, 2007; Reader, 1996, 2000).

The above accounts for the internal logic in this thesis addressing Japanese ritual events but there is also external motivation provided from the literature. Specifically, although field studies conducted with actual ritual communities remain extremely rare within CSR, one recent study conducted by Xygalatas et al. (2013b) provided compelling evidence that extreme ritual events (in this case, a Hindu festival in Mauritius that involved painful piercing) promote prosocial behaviour amongst both performers and observers, and result in the endorsement of more inclusive social identities. This study has been discussed in detail in Chapter 1 and will not be reviewed again here; however, the relevant point is that there have currently been no independent replications of the findings reported. One of the core motivations for this chapter then is to provide an independent conceptual replication and in so doing to address the extent to which the patterns reported appear to be cross-culturally recurrent.

Xygalatas et al.'s (2013b) findings concerning prosocial ingroup donations dovetail with predictions from Modes theory about elevated fusion and associated prosocial behaviour. However, in other respects the accounts diverge as Modes theory implies that the strongest relational bonds are forged through joint participation in high arousal ritual events and hence that performers should be the most prosocial group members, particularly when it

comes to making costly sacrifices (Swann, Gómez, Huici, et al., 2010; Swann et al., 2014). Costly signalling theories would seem to predict the inverse, that observers will be prompted more strongly to contribute to their ingroup after witnessing other members engage in costly physical displays of commitment (Bulbulia, 2004a; Henrich, 2009). Xygalatas et al.'s (2013b) study however found no significant difference between observers or performers, in relation to both prosociality and endorsement of inclusive identities although it did identify a non-significant trend for observers to donate more. The studies in this chapter seek to clarify these issues while simultaneously testing the predictions that have been examined in the previous chapters.

Field Study- Extreme Shinto Firewalking Festivals

Hypotheses

H1: Observers and performers of dysphoric fire walking rituals will display greater prosocial behaviour (directed at their ingroup) in the post-event period (Xygalatas et al. 2013b; Bastian et al. 2014; Bulbulia & Sosis, 2011).

H2: Observers and performers of dysphoric fire walking rituals will display more inclusive social identities post-event (Xygalatas et al. 2013b; Purzycki & Arakchaa, 2013, Bulbulia & Sosis, 2011; Haidt et al. 2008).

H3: Performers in dysphoric fire walking rituals will demonstrate higher levels of identity fusion with co-performers, and stronger *core bonds* with their ingroup post-event (Swann, Buhrmester et al. 2012), than observers.

H4: Higher levels of identity fusion with the ingroup will a) be associated with a greater willingness to perform costly sacrifices (Swann, Gómez, Huici, et al., 2010; Swann et al., 2014, 2012) and b) these relationships will be mediated by perceptions of *core bonds*.

H5: Those with more experience of firewalking performances will display higher levels of identity fusion with relational targets than other ingroup members, when controlling for confounding factors. In line with recent suggestions that dysphoric imagistic ritual experiences promote relational bonds with others who have shared the same self-defining experiences (Swann et al., 2012; Whitehouse & Lanman, 2014).

H6: The perception of firewalking experiences involving greater pain will be associated with greater prosocial contributions towards the ingroup (Bastian, Jetten, & Ferris, 2014; Olivola & Shafir, 2013).

4.3 Methods

Ethics

The design for the fieldwork study was primarily derived from the methodology reported by Xygalatas et al. (2013b). Questionnaires were designed by the thesis author and revised in collaboration with Masaki Yuki and his lab members at Hokkaido University. The current thesis author produced all necessary fieldwork forms, and managed a team of graduate students from Masaki Yuki's lab during data collection in the field. Translations of material into Japanese were produced by the study author, in consultation with two independent professional translators, and were then revised in lab meetings at the Culture, Social Ecology, and Psychology lab at Hokkaido University.

The Central University Research Ethics Committee (CUREC) reviewed this research and approval was received for the fieldwork protocol (SAME/CUREC1A/12-14). All procedures conformed to the ethical codes prescribed by Oxford University. Approval was also granted by the *Center for Experimental Research in Social Sciences* (CERSS) at Hokkaido University, who reviewed the fieldwork protocols and granted approval, under the supervision of Masaki Yuki.

A number of months prior to attendance at the field sites in Shakotan and Furubira, the relevant local shrine authorities responsible for organising the festivals were contacted in order to introduce the core research team, explain the goals of the study, and to seek permission for attendance. Copies of the questionnaire were provided in advance to the shrine authorities, upon their request, and permission was granted for data collection. During the collection periods all researchers wore identifying armbands and nametags and all questionnaires were identified with Hokkaido University and Oxford University logos.

All respondents approached in the field were provided with a brief summary of the questionnaire, the time required to complete, and the general purpose of the study without revealing any specific hypotheses. Those who agreed to participate were required to read and sign full consent forms, containing more detailed information about the study and relevant contact details of the thesis author. Upon completion of the survey all participants were compensated with 500 ¥ (approx. £3) for their time. For any participants who refused payment their allocation was donated to the shrine. The study did not involve any deception and the fieldwork conditions made any detailed debriefing at the study's goals

impractical. However, all respondents were provided with contact cards if they expressed any interest in receiving more information about the study.

Field sites

The two field sites visited for this study were *Bikuni Shrine* (美国神社) in Shakotan and *Kotohiro Shrine* (琴平神社) in Furubira, both located in the Shakotan peninsula of Hokkaido in Northern Japan. Shakotan and Furubira are neighbouring towns, located about 6km apart, and their primary industries are fishing and tourism. As of 2015, Furubira has a slightly larger population of around 3,751 residents, compared with Shakotan, which has 2,266⁶¹. Both towns hold dramatic firewalking festivals during the month of July, with firewalking performances conducted over the course of two nights.

The firewalking events held in these two towns are especially notable due to their inclusion of a masked performer, dressed as a creature from Japanese folklore called a *Tengu*⁶², and the size of the flames that the performers pass through. In both towns the celebrations follow a similar pattern; during the early morning and afternoon a local performer in an elaborate *tengu* costume accompanies groups of mostly younger men and women as they carry heavy portable wooden shrines (*mikoshi*) through the streets. Along the way, the parades stop occasionally so that the *tengu* performer may rest, while others in the parade perform short dance sequences for onlookers. Local residents provide offerings

⁶¹ Data available from <http://www.town.furubira.lg.jp/> and <http://town-shakotan.com/>

⁶² The original concept of the *tengu* is believed to have derived from the Chinese legendary concept of a 'heavenly dog' (天狗) and migrated to Japan from China, along with Buddhism, sometime during the 6-8th Centuries. Following their arrival in Japan, these 'heavenly dogs' underwent a dramatic transformation; morphing first, from celestial canines into monstrous demon birds, and then later, becoming anthropomorphised into long nosed, mischievous spirits, sometimes portrayed as lesser deities associated with mountain asceticism.

and donations which are collected by volunteers as the parades pass. In the evening, celebrations climax in a series of dramatic fire walking performances as the *tengu* performer followed by groups carrying the *mikoshi* and local boys in dragon costumes charge over and through the flames multiple times.

Figure 4-3: *Tengu* performer passing through the flames



Source: Image taken by author, used with permission

Before the *tengu*'s performance, water is drunk and blown over his costume by a shrine volunteer in order to help prevent the costume from catching fire. This often proves only partially successful and after each pass through the flames, the performer is patted down to extinguish flames emanating from the costume's beard or the single pronged wooden sandals he wears. For the *mikoshi* groups, the performance is accompanied by rhythmic jumping, the blowing of whistles, and repeated shouts of *wasshoi*. Some of the *mikoshi* performers are intoxicated and there are occasionally accidents in which members receive burns or in the worse cases fall into the fire. Most of the *mikoshi* performers have their faces

wrapped in damp cloth to prevent their face from being singed or inhaling too much smoke.

Figure 4-4: Mikoshi group passing through the flames



Source: Image taken by author, used with permission

The firewalking performances are usually held in the grounds of the local Shinto Shrine, which have a variety of temporary stalls set up to sell goods and snacks. However, in Furubira for the final day, the firewalking was performed in a local park that had been prepared with the necessary materials and festival stalls. The fires used in the performances are fuelled by bags of wood chip, which are constantly being replenished and stoked higher by local volunteers.

After the performances conclude, the grounds are quickly emptied as locals gather to celebrate in their homes or at local bars and visitors return to their cars or tour buses to make the trip home. The *mikoshi* are also transported back to the shrine or storage garages

by the relevant groups of performers. The firewalking performances last for about an hour in total.

Recruitment & Interviews

A group of four research students and the thesis author recruited all respondents in the study. Questionnaires were collected using paper & pencil methods from performers and observers located in and around the shrine grounds. Responses were collected during two periods: i) before the firewalking performances, and ii) immediately after. Locals had been informed in advance that the researchers would be present and performers had agreed to provide responses but nevertheless recruitment, especially in the post period, proved difficult. It was impossible to collect responses during the ritual due to the noise and the attention of all audience members being on the firewalking. Therefore, some post-event responses were collected in garages and in front of homes as this was the only way to contact performers who had left the grounds. Responses were collected on both days of the performances in Shakotan and Furubira. A series of unstructured interviews were also carried out in the months leading up to the event, and on the days of the performances, with the shrine priests, local organisers, residents, and visitors, via telephone, email, and face to face discussions, in order to gather logistic and general ethnographic information about the event.

Sample

A total of N=274 respondents, comprised of 55.3% (N=151) Male and 43.2% (N=118) Female⁶³, completed questionnaires across both sites. 37.7% (N=103) of the responses were collected in Shakotan and 62.3% (N=170) came from Furubira. All participants were Japanese nationals. The age range of respondents was from 11 to 81 years (M= 31.57, SD=17.46). Not all respondents answered every question and hence N is reported in the results section for each specific analysis. 21.8% (N=57) of the sample were firewalking performers and the remainder were comprised of observers 61.8% (N=162), vendors 6.1% (N=16) and volunteers 9.5% (N=26). 63.4% (N=151) of the sample were identified as locals and 36.6% (N=87) as outsiders.

Questionnaire Content

The full questionnaire is provided in Appendix VII but a summary of the main sections and details of the key measures included are provided below (Table 4.1). Overall, the questionnaire was designed to collect detailed information about people's response to dysphoric firewalking festivals, and to explore whether substantial differences in prosocial attitudes and behaviour, inclusive self-identification, and levels of identity fusion were detectable i) in pre- and post- performance samples and ii) between performers and observers. Potentially confounding demographic information was collected, along with details about previous involvement with the festivals. In particular, care was taken to establish whether individuals were locals or external visitors. This was distinguished

⁶³ 1.5% (N=4) were missing responses.

⁶⁴ All respondents under 15 completed the questionnaire only after consent was obtained from a guardian. Compensation payments were provided to the guardian.

through self-selection of specific roles (local/outside observer and local/outside performer) and examination of demographic details (location of hometown, time spent living in Furubira/Shakotan). Perceptions of pain involved were also collected in order to examine predictions made by costly signalling theories.

Identity fusion was measured with different targets, including firewalking performers, the local community, and Japan as a whole, using the validated visual identity fusion measure (Swann, Gómez, Huici, et al., 2010; Swann et al., 2009). The visual measure was chosen over the verbal measure to increase ease of understanding and reduce the amount of time required. Different targets were used in order to reduce demand characteristics and enable an examination of whether any effects of the performance were directed at relational or broader categorical groups.

Prosociality was operationalised through three measures; two self-report items and a behavioural voluntary donation task. The two self-report items were based on hypothetical scenarios and addressed: 1) willingness to donate time to the local shrine, and 2) willingness to risk bodily harm to aid a firewalking performer. The first item was measured using hours donated as an outcome, while the second item was measured on a 5-point scale ranging from 1- *I definitely would not act and avoid any risk* to 5- *I definitely would act, despite the risk*. The behavioural measure replicated the design used in the BJJ study (Chapter 2) and asked respondents, who voluntarily choose to participate in a lottery to win 15,000 ¥ (approx. £100), to designate if they would like to donate any of their reward to the shrine if they were selected as a winner. Respondents were informed that they could keep all of the money, donate all of it to their shrine, or split the amount in any ratio they preferred and

were reassured that any amount donated would be sent anonymously, to remove potential status-seeking behaviour. Both this donation task and self-reported willingness to risk bodily harm were also designated as costly sacrifice measures and it was anticipated that they would display stronger associations with fusion than the less severe time donation measure.

The self-identification measure was adapted from (Xygalatas et al., 2013b), and involved respondents choosing from a superordinate inclusive national identity (Japanese) and a more exclusive local identity (Shakotan/Furubira Person)⁶⁵. The original measure contrasted a national identity with a religious identity but for the current study regional identity was selected as an alternative anchor as religious identities are not exclusive or salient in Japan. As with the original study, the measure was presented as a horizontal line of 18cm that respondents were asked to place a mark on to indicate where they felt their identity was best represented 'right now'. Responses were coded according to 0.5cm intervals resulting in a 36-point scale. A brief summary of the questionnaire content is provided on the following page.

⁶⁵ The expression used in Japanese sounds more natural than the English translation.

Table 4-1: Shinto firewalking questionnaire content

	SECTION	SAMPLE QUESTIONS & SCALES	MEASURE
1	General Demographic Information	Sex, Age, Nationality, Marital Status, Occupation	Various
		Current Residence & Duration of Residence	Years
2	Social Identity	Inclusive/ Exclusive identity measure (Xygalatas et al., 2013b) Adapted to Japanese context using local town (exclusive) vs. Japanese (inclusive) poles.	18-point scale
3	Mood	Affective Valence (Bradley & Lang, 1994) – Using Self-Assessment Manikin	9-point scale
		Activation (Bradley & Lang, 1994) – Using Self-Assessment Manikin	9-point scale
4	Identity Fusion	Pictorial measure of identity fusion (Swann et al., 2009): - with <i>Tengu</i> performer - with <i>Mikoshi</i> performers - with Shakotan/Furubira Community - with Japan	7-point scale
5	Prosocial Behaviour	Scenario 1: Prosocial behaviour <i>If the shrine requested the community to help advertise/promote the matsuri, how many hours during a holiday week would you be willing to donate to take part in group activities aimed at promoting the matsuri? Please answer in hours, using only numbers e.g. 10 = willing to donate 10 hours over one holiday week.</i>	Hours
		Scenario 2: Costly sacrifice <i>During the firewalking performance you recognise that a performer is likely to fall and sustain a severe burn. You are the only person in a position to help but in doing so would place yourself at great risk to sustain a severe burn instead. How likely would you be to act to intervene?</i>	5-point scale
6	Previous Festival Experience	Previous Festivals Attended	Festivals
		Time since last attendance	Years
		Previous experiences as firewalker	Times
		Time since last firewalking performance	Years
7	Current Festival Experience	Role in this year's festival	Category
		Close friends in firewalking performances	Count
		Family members in firewalking performances	Count
		Perceived pain of: - <i>Tengu</i> performer - <i>Mikoshi</i> performers	5-point scale
8	Subjective Response to Festival	2-items measuring perception of festival as a sacred value (Fischer et al., 2013)	7-point scale
		2-item measure of personal consequentiality: <i>Participating in this matsuri has helped shape the person I am today.</i> <i>This matsuri is very important for me personally.</i>	7-point scale
		1-item measure of community importance: <i>This matsuri is very important for the community.</i>	7-point scale
		3-item measure of relational bonds: <i>I feel an emotional attachment to participants of the matsuri.</i> <i>I feel like I share something meaningful with participants in the matsuri.</i> <i>The firewalking performers are like family to me. (Buhrmester et al., 2014)</i>	7-point scale
9	Prize Lottery	Voluntary donation opportunity to shrine	0-15,000 円

4.4 Results

Data Preparation

Subjective Response to Festival Items

- Sacred Values

Two items measuring perception of festival as a sacred value from Fischer et al. (2013) were collected using a 7-point scale. The motivation for including this measure was due to recent findings that it served as a moderating variable in the relationship between fusion and willingness to make costly sacrifices. The status measures displayed a weak correlation between items ($r=.360$, $p<.001$) and hence they could not be combined into a viable scale ($\alpha=.529$). As a result they were included in a principal components analysis reported below.

- Principal Components Analysis (Core Bonds)

Eight items were used to measure subjective responses to the firewalking festival and these were anticipated to address the following four concepts: Two items on perception of the festival as a sacred value (Fischer et al., 2013); one-item on the importance of the event for the community; three items on the relational bonds felt with other festival participants (see Chapters 2 & 3); and two items on personal consequentiality. Five of the measures were not taken from validated scales and all items had been translated into Japanese hence an exploratory principal components analysis was conducted on all eight items.

Due to the nature of field based data collection, some respondents provide low quality responses. To partially control for this, standard deviations on the responses provided to the eight items mentioned above were calculated. Responses which revealed low deviation

(e.g. $SD < .50$) ($N=39$) and those with incomplete responses ($N=13$) were excluded from the PCA and all subsequent analyses based on these items.

Table 4-2: Rotated factor loadings for response to festival items

Item	Rotated Factor Loadings	
	Core Bonds	Sacred Values
Shared Meaning	.866	-
Self-Transformational	.864	-
Emotional Attachment	.793	-
Personal Importance	.731	-
Family Tie	.687	-
Sacred Value 1 (Beyond Money)	-	.861
Sacred Value 2 (Never Abandon)	-	.727
Eigenvalues	3.753	1.000
% of variance	53.61	14.28
Cronbach's α	.876	.529

Component loadings over 0.6 are highlighted. Component loadings < 0.4 not shown.

The suitability of a PCA was assessed prior to analysis and an inspection of the correlation matrix showed that all variables had at least one correlation coefficient greater than .3. The *community importance* item had a large number of similar sized correlations and no individual correlation greater than .5 and was thus excluded from further analysis. The overall Kaiser-Meyer-Olkin (KMO) measure of the remaining 7 items was .847, with individual KMO measures all greater than .8, classifications of 'meritorious' according to Kaiser & Rice (1974). Furthermore, Bartlett's test of sphericity was statistically significant ($p < .001$), indicating that the data were likely factorizable.

The resulting PCA revealed two components that had eigenvalues greater than one, which explained 53.61% and 14.29% of the total variance, respectively. Visual inspection of the scree plot indicated that two components should be retained (Cattell, 1966). The two-component solution produced suggested that the relational bonds and personal

consequentiality items should be treated as one factor and the sacred value items as another. This met the interpretability criterion and, as such, two components were retained for the final analysis. The two-component solution explained 67.90% of the total variance with a Varimax orthogonal rotation being employed to aid interpretability.

The combination of personal consequentiality and relational bonds items on Component 1 were labelled as representing perceptions of *Core Bonds* as the measure encapsulated not just aspects concerning social affiliation but also indications of how central the experience was to an individual's identity. Component 2 was more straightforward to interpret as sacred values. Component loadings and communalities of the rotated solution are presented in Table 4.2. *Core Bonds* items displayed excellent reliability when combined into a single scale. The two sacred value items did not form a reliable scale, as discussed above, and are treated separately in the results reported below.

Hypothesis 1 (Post Event= More Prosocial)

Responses from observers and firewalking performers were compared independently in the pre- and post-event periods on three measures of prosociality: two self-report and one behavioural. Two response groups were excluded from analyses volunteers and vendors. The justification for this is that there were only six responses in the pre-event period for volunteers and although they did not participate in the firewalking they were also not passive observers of the event, making them difficult to classify. Conversely, vendors were excluded since they were unable to observe performances and, in general, displayed little interest in the festival.

Table 4-3: Descriptive statistics for H1 variables

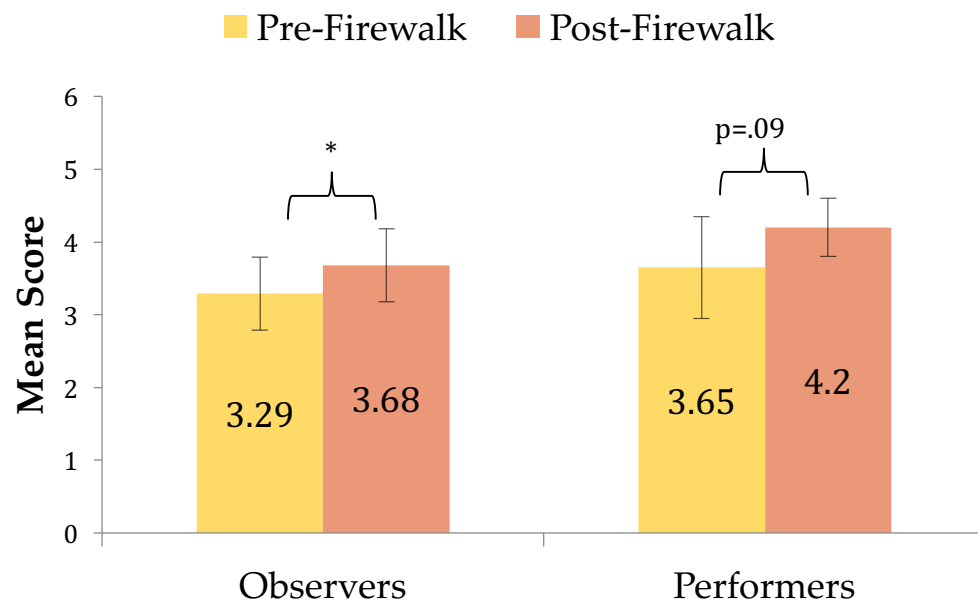
Variable	Skew (SE)	Kurt. (SE)	M (SD)
Donate Time	2.382 (.162)	7.651 (.332)	13.24 (16.198)
Risk Injury	-.391 (.152)	-.199 (.316)	3.54 (1.078)
Donation to Shrine	.075 (.199)	-1.132 (.396)	1695 (9801)

The donate time outcome displayed a severe non-parametric distribution (Table 4.3), due to 11 extreme responses. However, the responses were from valid questionnaires so they were retained in analysis using non-parametric Mann-Whitney U tests. Comparing the two time periods, median time donations were found to not be significantly different for observers, $N=151, U=2,710, z=.964, p=.335$, or performers, $N=53, U=321.500, z=-.485, p=.628$, using an exact sampling distribution for U (Dinneen & Blakesley, 1973).

For willingness to risk injury, a two-tailed independent t-test identified a significant increase, $t(153)=2.199, p=.029, d=0.355$, in stated willingness amongst observers from the pre- ($N=107, M=3.29, SD=1.046$) to post-event ($N=48, M=3.68, SD=.937$) period. However, when local and non-local observers (see questionnaire content section for details) were examined separately no significant change was found for either group: Local Observers- $t(68)=1.179, p=.243$; Non-Local Observers- $t(83)=1.535, p=.129$.

The same pattern of higher values in the post-event period was also represented amongst firewalkers (Pre-Firewalk, $N=26, M=3.65, SD=1.413$; Post-Firewalk, $N=30, M=4.20, SD=.887$), but the difference was found to be statistically insignificant, $t(40.880)=1.702, p=.096, d=0.478$. Levene's test indicated unequal variances, $F=5.686, p=.021$, so degrees of freedom were adjusted accordingly. The mean difference in willingness to risk injury across the periods was .546, however 95% CI [-.102, 1.194] crossed 0.

Figure 4-5: Willingness to risk injury in pre/post periods



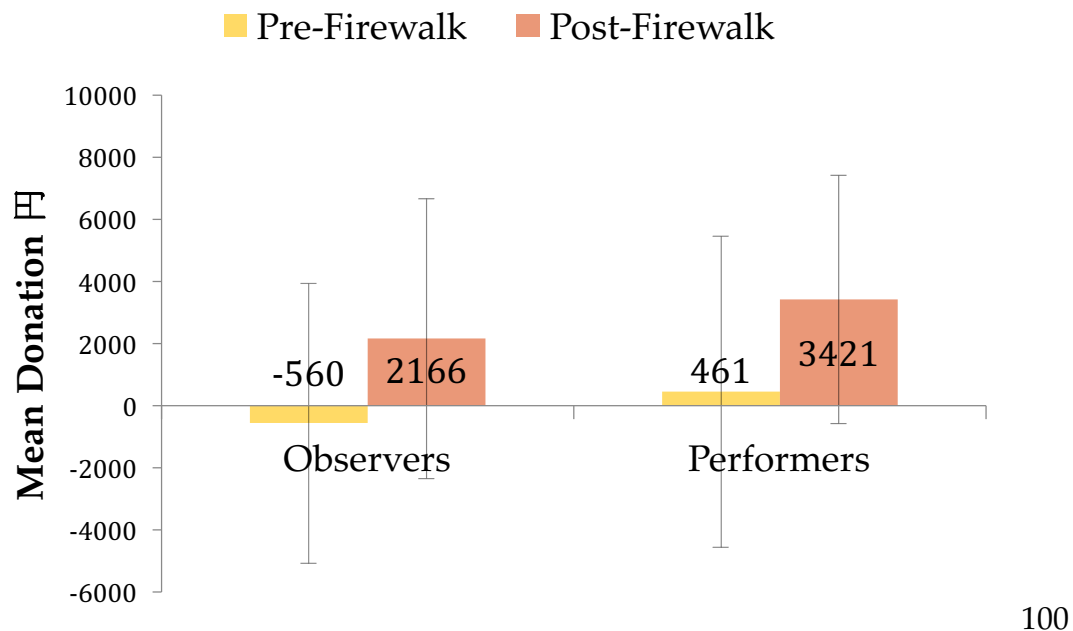
Note: Error bars represent standard deviations.

Finally, for the behavioural donation measure, visual inspection of histograms and Shapiro-Wilk's tests of normality indicated significantly non-parametric distributions. In particular, there were multimodal peaks in the data created by 34.6% ($N=66$) of respondents choosing to donate all of their potential winnings 15,000 ¥, 16.7% ($N=31$) donating 10,000 ¥, and 31.7% ($N=56$) donating 5000 ¥, which generated large standard deviations. Converting the responses into ordinal categories of low donations (0-5000 ¥), medium (5001-10,000 ¥), and high (10,001 to 15,000 ¥), reduced the magnitude of deviations but did not result in any outcomes different from those reported below.

Using non-parametric Mann-Whitney U tests, no significant difference, $N=104$, $U=1,213$, $z=1.136$, $p=.256$, was found between median donations allocated by observers in the pre- and post-event periods. Similarly, there was no significant difference, $N=34$, $U=165.5$, $z=1.059$, $p=.289$, found in the median donations allocated by performers. Visually

there did appear to be evidence for a positive trend in the data in favour of the post-event period but wide standard deviations make the trends unreliable.

Figure 4-6: Monetary donations in pre/post periods



Mean Donations (Y-axis) calculated by subtracting personal prize allocation from amount allocated to the shrine, Note: Error bars represent standard deviations.

In summary, significant increases in prosociality were only detected on one of the self-report measures. Contrary to predictions, both voluntary donations and stated willingness to donate time displayed no significant increase for either performers or observers. For the donation measure a clear limitation to the validity of the findings is the low variation in distributions of responses, with most participants choosing to donate in multiple of 5,000 ¥. The original measure was anticipated to range between 0-15,000 ¥ but the data suggests that the measure was more commonly responded to as a four-point scale (0, 5000 ¥, 10000 ¥, 15000 ¥).

An increase in willingness to risk injury to assist a festival performer was found for both performers and observers in the post-event period. Furthermore, this outcome measure

displayed the most normally distributed responses, enabling parametric comparisons to be conducted. However, the difference observed was only found to be close to significance for performers, and for observers the increase in values was no longer significant when local and external observers were examined independently. This suggests that the difference found for observers may simply be an artefact of there being more non-local observers in the post event period.

Hypothesis 2: (Post-Event = More Inclusive Social Identity)

Responses from observers and firewalking performers were compared independently in the pre- and post-event periods for endorsement of inclusive (national) vs exclusive (local) social identities, measured on a single item 18-point scale. As with H1, volunteers and vendors were excluded from analysis.

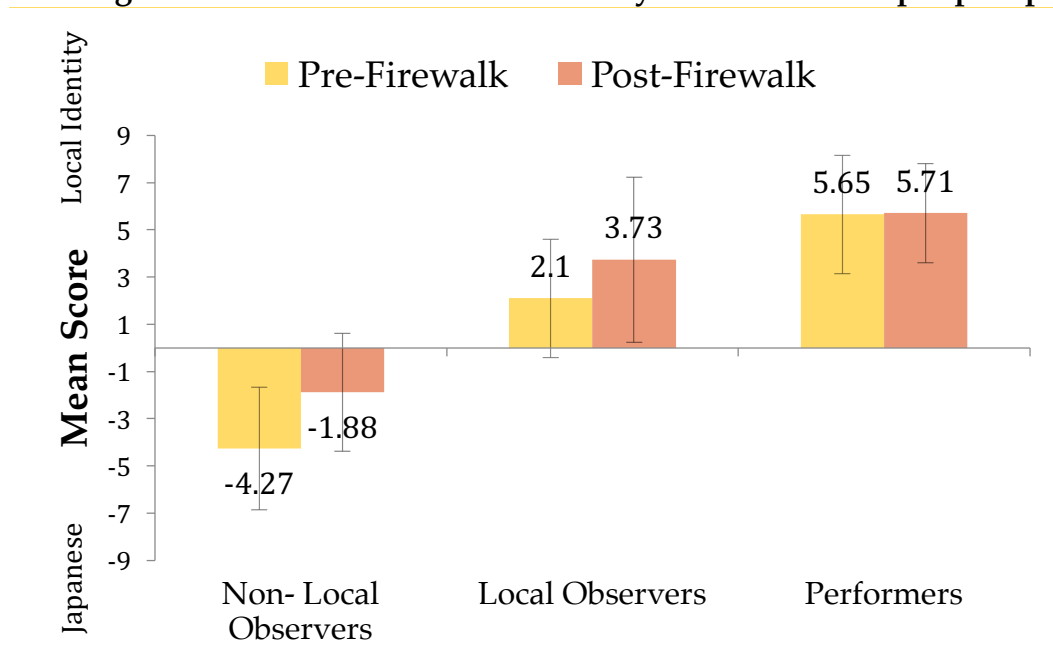
Table 4-4: Descriptive statistics for H2 variables

Variable	Skew (SE)	Kurt. (SE)	M (SD)
Ident. (Performers)	-1.721 (.333)	2.227 (.656)	14.686 (4.627)
Ident. (Ext. Observers)	.922 (.274)	-.418 (.541)	5.104 (5.401)
Ident (Local Observers)	-.634 (.295)	.769 (.582)	11.841 (5.159)

Distributions of the identification measure were non-normal for both observers and firewalkers, as assessed by visual inspection of histograms and significant Shapiro-Wilk tests. For observers, this was primarily due to bipolar distribution caused by a positively skewed distribution (towards national identity) amongst non-local observers, and a negatively skewed distribution (towards local identity) by local observers. For performers it was the result of a substantial negative skew. Therefore, all of these groups were examined independently, using non-parametric Mann-Whitney U tests.

No significant difference was found in median identification between time periods for local observers, $N=66$, $U=652$, $z=1.451$, $p=.147$, non-local observers, $N=77$, $U=476.500$, $z=1.221$, $p=.222$, and performers, $N=51$, $U=277.5$, $z=-.863$, $p=.388$, using an exact sampling distribution for U (Dinneen & Blakesley, 1973). Performers show no indication of change but that may have been due to a high baseline measure in the pre-event period.

Figure 4-7: Inclusive/Exclusive identity endorsement in pre/post periods



Note: Error bars represent standard deviations.

Independent comparisons between the three groups' endorsement of inclusive identities in the pre- and post-event periods were conducted using Kruskal-Wallis tests. As discussed, the distribution of scores was not similar for all groups and therefore comparisons of medians were not conducted. Identification scores were found to be significantly different between the groups in both the pre-event, $N=124$, $\chi^2(2)=49.814$, $p<.001$, and post event periods, $N=70$, $\chi^2(2)=12.264$, $p=.002$. These differences were further explored using pairwise

comparisons, performed using Dunn's (1964) procedure, applying a Bonferroni correction for multiple comparisons.

In the pre-event period, the pairwise comparisons revealed significant differences ($p < .001$) in identification scores between performers (Mean rank=98.26) and non-local observers (Mean rank=41.97) and between local (Mean rank=76.72) and non-local observers ($p < .001$). However, there was only a marginal difference between local observers and performers when the adjusted significance value was applied ($p = 0.07$). In the post event period, a similar pattern was observed, again there was a significant difference ($p = .002$) in identification scores between performers (Mean rank=43.57) and outside observers (Mean rank=19.38) and a difference between local (Mean rank=34.42) and outside observers that was not statistically significant ($p = .087$) using standard cut off values and adjusted p values. Furthermore, there was no significant difference ($p = .251$) detected between performers' and local observers' levels of identification. These findings accord with expectations but they provide a clear illustration of why it is essential to distinguish between different types of 'observers'.

In summary, there was no evidence for an increase in inclusive self-identification in the post-event period when the firewalking performance was complete. Conversely, non-significant trends were found amongst observers to endorse more exclusive/local identities. Counter to the findings of Xygalatas et al. (2013b), performers and local observers endorsed more exclusive identities in both periods and only outside observers endorsed a stronger inclusive identity.

Hypothesis 3 (Performers = More Fusion & Core Bonds with Relational Targets)

Responses from observers and firewalking performers were compared independently in the pre- and post-event periods on the visual identity fusion measure with four separate targets and *core bonds*. Volunteers and vendors were excluded from analyses, along with 10 cases that showed very low deviation ($SD < .5$) in responses on fusion items and the eight items in the subjective response to festival section.

Table 4-5: Descriptive statistics for H3 variables

Variable	Skew (SE)	Kurt. (SE)	M (SD)
Visual Fusion (Tengu)	1.218 (.168)	.395 (.334)	2.56 (1.858)
Visual Fusion (Firewalkers)	.737 (.168)	-.901 (.334)	3.13 (2.149)
Visual Fusion (Community)	.193 (.169)	-1.395 (.336)	3.89 (2.183)
Visual Fusion (Japan)	-.049 (.169)	-1.509 (.333)	4.41 (2.214)
Core Bonds (Japan)	-.742 (.167)	.219 (.333)	5.34 (1.379)

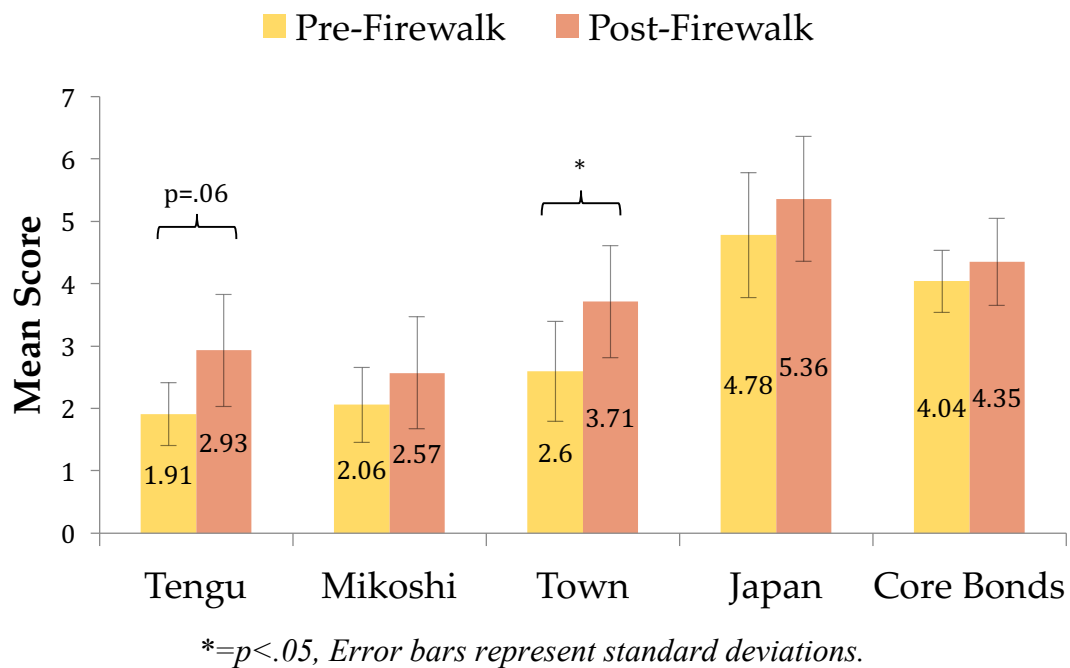
Distributions of the fusion measures were non-normal for both observers and firewalkers, as assessed by visual inspection of histograms and significant Shapiro-Wilk tests. Therefore, all of these groups were examined independently, using non-parametric Mann-Whitney U tests.

Focusing first on non-local observers ($N=84$), there were no significant differences found for fusion with the *mikoshi* performers ($p=.514$) or with Japan ($p=.297$). However, there was an increase (Pre-event- Mean rank=40.41; Post event- Mean rank=52.96) in fusion with the *Tengu* performer that was just outside of statistical significance, $N=84$, $U=636.500$, $z=1.870$, $p=.061$, and a significant increase with the local community, $N=81$, $U=640$, $z=2.189$, $p=.029$.

These results must be treated cautiously due to the small sample ($N=14$) in the post event period. Most non-local observers left as soon as the festival had ended which meant

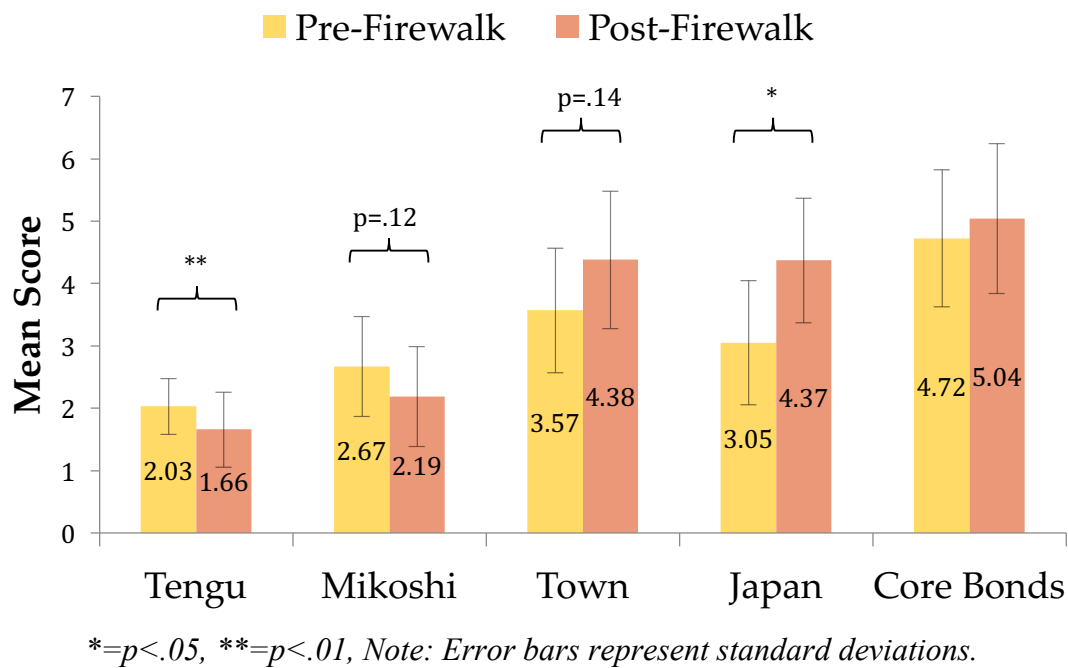
that the 14 responses collected may have been from individuals who identified as being from out of town but had some existing local ties. For the *core bonds* measure there was also no significant difference between time periods ($p=.334$).

Figure 4-8: Fusion & Bonds in pre-post periods for non-local observers



Turning to local observers ($N=69$), there were non-significant trends found for a decrease in fusion with the *mikoshi* performers ($p=.120$) and an increase with the local community ($p=.142$). There was a significant decrease (Pre-event- Mean rank=40.16; Post event- Mean rank=29.03) in fusion with the *Tengu* performer, $N=69$, $U=401$, $z=-2.483$, $p=.013$, and a significant increase with fusion with Japan, $N=67$, $U=707.500$, $z=1.964$, $p=.050$. For the *core bonds* measure, there was no significant difference between time periods ($p=.334$). Additionally, the relatively balanced samples in the pre ($N=30$) and post event ($N=37$) period, increases confidence in the validity of the findings.

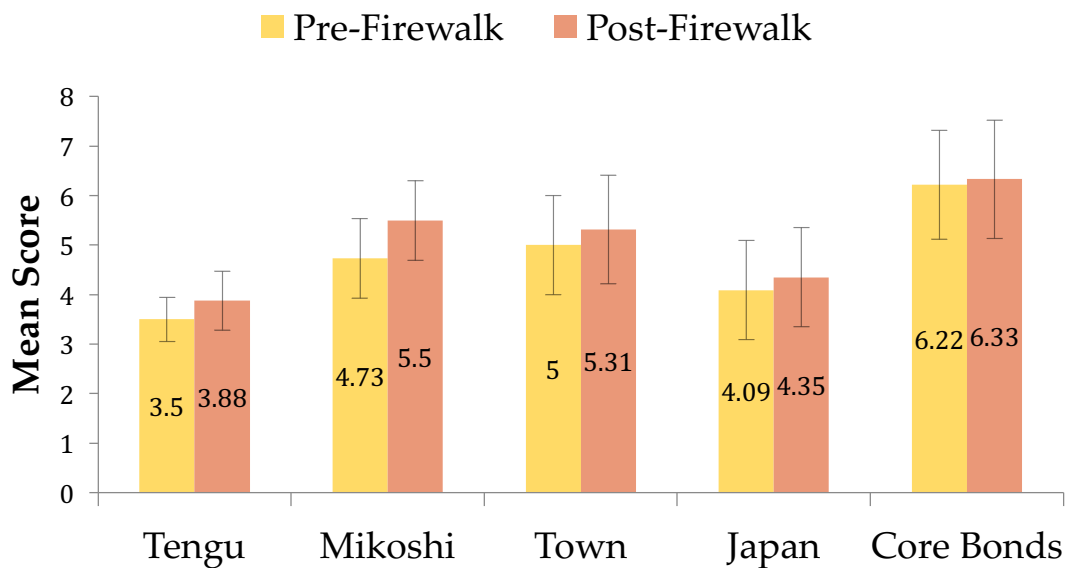
Figure 4-9: Fusion & Bonds in pre-post periods for local observers



Finally, examining the responses from the *mikoshi* performers ($N=69$), there were no significant changes in any of the fusion measures or the core bond measure (all $p > .247$). This seems likely to represent a ceiling effect as their pre-event baseline measures were already elevated when compared with the observer's responses.

The differences in fusion and *core bonds* were examined between groups in the critical post-event period, using Kruskal-Wallis tests. The distribution of scores were not similar for all groups and therefore comparisons of medians were not conducted. Fusion scores were found to be significantly different between the groups for the *Tengu* performer, $N=72$, $\chi^2(2)=19.617$, $p < .001$; *Mikoshi* performers, $N=72$, $\chi^2(2)=32.725$, $p < .001$; and *core bonds*, $N=73$, $\chi^2(2)=20.973$, $p < .001$. There was also a marginally significant difference in fusion with local community, $N=72$, $\chi^2(2)=5.850$, $p=.054$, but no significant difference between fusion with country, $N=70$, $\chi^2(2)=2.363$, $p=.307$. Finally, *core bonds* also differed significantly between groups, $N=73$, $\chi^2(2)=20.973$, $p < .001$.

Figure 4-10: Fusion & Bonds in pre-post periods for performers



**= $p < .05$, Note: Error bars represent standard deviations.*

Significant main differences were further examined with pairwise comparisons, performed using Dunn's (1964) procedure, applying a Bonferroni correction for multiple comparisons. These revealed that the differences detected in fusion with the *Tengu* performer was being driven ($p < .001$) by a comparison between local observers and performers (Mean rank=48.58), with local observers having a significantly lower level (Mean rank=25.36). For the difference observed in fusion levels with the *mikoshi* performers, the effect was being driven by the elevated fusion of the performers (Mean rank=54.87) which was significantly higher ($p < .001$) than both local (Mean rank=25.14) and outside observers (Mean rank=28.36). There was no significant difference observed between both types of observers ($p = .625$). This pattern was also replicated with the *core bonds* measure, as performers displayed a significantly ($p < .001$) elevated distribution (Mean rank=51.46) when compared with both local (Mean rank=31.75) and outside observers (Mean rank=21.81).

In summary, while performers displayed elevated levels of fusion with all other target performers in the post event period, there was no significant increase in these values between the pre- and post-event periods. High baseline scores, potentially representing ceiling effects, seem likely to have played a role in this as even in the pre-event period performers displayed elevated levels of fusion when compared with observers. Local observers unexpectedly demonstrated lower levels of fusion with performers in the post-event period and, instead, higher levels of fusion were found marginally with local community and, significantly, with Japan. Non-local observers conversely showed no increase in fusion with Japan, but higher levels of fusion with the local community and the *Tengu* performer.

Hypothesis 4 (Higher Fusion= More Costly Sacrifice)

Using the full sample collected, excluding vendors and low quality responses, the role of fusion in promoting prosocial behaviour and costly sacrifices was examined through a simple spearman's correlation analysis (Table 4.6).

Table 4-6: Spearman's correlation matrix for fusion and sacrifice measures

Variable	Willingness to Donate Time	Willingness to Risk Injury	Voluntary Prize Donation
Tengu Fusion	.271***	.219***	.208*
Mikoshi Fusion	.265***	.247***	.206*
Town Fusion	.243***	.335***	.005
Japan Fusion	.207***	.209***	.133
Core Bond	.401***	.381***	.201

$N=223$, *= $p<.05$, **= $p<.01$, *** = $p<.001$

Correlations of a similar magnitude were found with all fusion scores for the two self-report measures; however, for the behavioural donation measure only fusion with the two performer categories emerged as significant. To examine the respective individual

contributions to predicting the sacrifice measures a series of multiple linear regressions were conducted on the three outcomes.

For all models, assumptions were tested in full but no indication of multicollinearity as assessed through correlation analysis and VIF and tolerance statistics was found. With the 14 low quality exclusions already applied standardized residuals identified no cases as extreme outliers. Finally, a visual examination of the resulting residuals in a histogram, P-P plot and Normal Q-Q plot demonstrated a normal distribution.

Table 4-7: Multiple regression on costly sacrifice measures

Variable	Outcome: Donate Time		Outcome: Risk Injury		Outcome: Donate Money	
	MODEL 1		MODEL 2		MODEL 3	
	B (SE)	β	B (SE)	β	B (SE)	β
Constant	-1.829 (2.933)	-	2.458 (.186)	-	-3546(2287)	-
Tengu.Fusion	1.600 (.699)	.177*	.062 (.044)	.105	742 (544)	.135
Mikoshi.Fusion	.914 (.635)	.115	.064 (.040)	.123	777 (495)	.162
Town.Fusion	.469 (.563)	.062	.100 (.036)	.201**	-585 (438)	-.128
Japan.Fusion	1.400 (.500)	.191**	.071 (.032)	.149*	671 (389)	.152*
R^2	.132		.183		.073	
F	7.894***		6.551***		2.634*	
N	211		219		137	

*= $p < .05$, **= $p < .01$, ***= $p < .001$, shading added to emphasis variables with significant coefficients.

The results suggested that for the willingness to donate time to the shrine during a crisis the strongest predictor was fusion with Japan and fusion with *Tengu* performer. Alternatively, for the willingness to risk injury measure, although fusion with Japan remained a significant predictor, fusion with local community was the strongest factor. Finally, in the behavioural measure and counter to initial correlation analyses, fusion with Japan emerged as the strongest predictor.

In summary, fusion with Japan was consistently the strongest predictor for costly sacrifices but for the more visceral risking injury item, fusion with local community was more strongly predictive. Surprisingly, fusion with *mikoshi* performers did not correlate with any of the outcome measures. In order to explore these associations better hierarchical regressions with other relevant identified confounds were conducted and are reported below.

H4 Sacrifice Measure 1 (Willingness to Donate Time)

To identify potentially confounding external variables, an initial exploratory multiple regression was conducted on willingness to donate time entering the following variables as predictors: demographic factors- age, sex, residence time, town; positive and negative mood; times attended festival; times performed as a firewalker; two sacred values measures; importance of festival to community and presence of family or friend in firewalkers, coded a binary measure.

Testing assumptions: there was acceptable independence of residuals, as assessed by a Durbin-Watson statistic of 1.926. There was also no indication of multicollinearity as assessed through correlation analysis and VIF and tolerance statistics. However, an examination of standardized residuals from the final model indicated 12 cases as extreme outliers with values over 3 SDs and a further 2 cases demonstrated extreme values of leverage and Mahalanobis distance. These 14 cases were removed and the model was rerun. Finally, a visual examination of the resulting residuals in a histogram, P-P plot and Normal Q-Q plot demonstrated a normal distribution.

The model that emerged significantly predicted willingness to donate time; $F(14,154)=2.957$, $p=.001$, $R^2=.212$. However only perception of *Tengu* performer's pain and total firewalking events performed emerged as significant predictors (Table 4.8). Consequently, these were the only variables retained as additional predictors for the subsequent hierarchical regression models.

Table 4-8: Exploratory regression with willingness to donate time as outcome

Variable	b (SE)	Standardized β	Sig.
Constant	-4.538 (9.307)	-	.609
Sex	-.292 (1.754)	-.013	.868
Age	.022 (.058)	.032	.711
Residence (Yrs)	.021 (.078)	.026	.786
Town	2.011 (1.932)	.084	.300
Tengu (Pain)	-2.242 (.904)	-.269*	.014
Mikoshi (Pain)	.428 (.989)	.050	.665
Sacred Value 1	.332 (.746)	.041	.657
Sacred Value 2	.679 (.719)	.083	.347
Community Imp.	1.501 (1.094)	.127	.172
Total Firewalk	.607 (.199)	.239**	.003
Total Festival	-.095 (.097)	.093	.325
Affect (Mood)	.441 (.667)	.054	.509
Activation (Arousal)	-.493 (.360)	-.112	.172
Relative Performer	1.285 (2.271)	.044	.572

$p<.05^*$, $p<.01^{**}$, shading added to emphasis significant predictors in the model

Based on the exploratory analysis, a three-stage hierarchical regression, with pair-wise deletion, was conducted on willingness to donate time with perceived *Tengu* pain and total firewalking performances entered as control variables in the first stage, all four identity fusion measures in the second, and finally adding *core bonds* for the final stage. *Core bonds* was retained for the final stage to enable potential mediating relationships with fusion measures to be identified. All assumptions were re-examined prior to analysis and an additional extreme 10 outliers (values greater than 3 SD) were excluded.

The model in the first stage significantly predicted willingness to donate time; $F(2,180)=12.286, p<.001, R^2=.120$, with perceived *Tengu* pain as the strongest contributor to the model. However, when the identity fusion measures were added in the second stage the model that emerged predicted a greater amount variance, $\Delta R^2=.183, F(4,176)=3.361, p=.011$, but only fusion with Japan appeared to contribute to the predictive power of the model. In the final stage, *core bonds* emerged as the strongest predictor (Table 4.9).

Table 4-9: Hierarchical regression on willingness to donate time

Variable	Outcome: Donate Time					
	STAGE 1		STAGE 2		STAGE 3	
	B (SE)	β	B (SE)	β	B (SE)	β
Constant	10.811 (2.518)	-	6.656 (1.657)	-	1.350 (2.091)	-
Tengu (Pain)	-1.060 (.347)	-.215**	-1.001 (.348)	-.203**	-.863 (.336)	-.175**
Firewalk (Total)	.385 (.109)	.249*	.280 (.114)	.181*	.239 (.110)	.155*
Tengu.Fusion	-	-	.422 (.320)	.114	.183 (.314)	.049
Mikoshi.Fusion	-	-	.310 (.302)	.094	.037 (.299)	.011
Town.Fusion	-	-	.056 (.254)	.018	-.270 (.258)	-.085
Japan.Fusion	-	-	.444 (.227)	.143*	.511 (.219)	.164*
Core Bonds	-	-	-	-	1.473 (.376)	.333***
R^2	.120		.183		.248	
F	12.286***		6.551***		8.266***	

$N=182$, $*=p<.05$, $**=p<.01$, $***=p<.001$, shading added to emphasis variables with significant coefficients.

These results indicated that only extended fusion was a significant predictor of willingness to donate time, when controlling for confounding factors, and that perceptions of *core bonds* appeared to be an independent and substantially stronger predictor.

H4 Sacrifice Measure 2 (Willingness to Risk Injury)

The same potentially confounding variables identified above were entered into a second exploratory regression to identify potentially confounding external variables on the willingness to risk injury self-report measure. As this outcome was intended to serve as a

more visceral costly sacrifice measure, it was anticipated that the fusion items would be stronger predictors of this outcome. Testing assumptions of the exploratory model: there was acceptable independence of residuals, as assessed by a Durbin-Watson statistic of 1.946. There was also no indication of multicollinearity as assessed through correlation analysis and VIF and tolerance statistics. After retaining the same 14 cases identified in the previous analysis as exclusions, no additional outliers were detected. Finally, a visual examination of the resulting residuals with histograms, P-P plots and Normal Q-Q plots demonstrated a normal distribution.

Table 4-10: Exploratory regression with willingness to risk injury as outcome

Variable	b (SE)	Standardized β	Sig.
Constant	2.844 (.814)	-	.001
Sex	-.113 (.161)	-.054	.485
Age	-.006 (.006)	-.096	.287
Residence (Yrs)	.006 (.007)	.078	.424
Town	-.083 (.177)	-.039	.642
Tengu (Pain)	-.020 (.083)	-.026	.810
Mikoshi (Pain)	.034 (.089)	.043	.704
Sacred Value 1	.135 (.069)	.188*	.050
Sacred Value 2	.048 (.064)	.068	.454
Community Imp.	.017 (.095)	.016	.862
Total Firewalk	.021 (.017)	.100	.221
Total Festival	.013 (.009)	.142	.142
Affect (Mood)	.023 (.061)	.030	.711
Activation (Arousal)	-.055 (.033)	-.136	.103
Relative Performer	-.059 (.214)	-.022	.572

p<.05, p<.01**, shading added to emphasis significant predictors in the model*

The model that emerged significantly predicted willingness to risk injury; $F(14,158)=2.294$, $p=.007$, $R^2=.169$ but only the first sacred value item emerged as significant predictors (Table 4.10). Although it did not reach the cut off value for statistical significance activation was also retained as a control item.

Based on the exploratory analysis, a three-stage hierarchical regression was conducted on willingness to risk injury with the sacred value measure and activation entered as control variables in the first stage, all four identity fusion measures in the second and *core bonds* added in the final stage. As before, the *core bonds* measure was retained for the final stage to enable potential mediating relationships with fusion measures to be identified. All assumptions were re-examined prior to analysis but no additional outliers (values greater than 3 SD) were identified.

Table 4-11: Hierarchical regression on willingness to risk injury

Variable	Outcome: Risk Injury					
	STAGE 1		STAGE 2		STAGE 3	
	B (SE)	β	B (SE)	β	B (SE)	β
Constant	2.548 (.340)	-	1.779 (.357)	-	1.678 (.361)	-
Sacred Value 1	.196 (.047)	.271***	.152 (.046)	.211***	.103 (.054)	.142
Activation	-.049 (.027)	-.122*	-.021 (.026)	.051	-.018 (.026)	-.045
Tengu.Fusion	-	-	.052 (.044)	.088	.043 (.045)	.073
Mikoshi.Fusion	-	-	.044 (.040)	.085	.030 (.041)	.723
Town.Fusion	-	-	.090 (.035)	.180*	.069 (.037)	.139
Japan.Fusion	-	-	.066 (.031)	.138*	.072 (.031)	.150*
Core Bonds	-	-	-	-	.102 (.061)	.148
R^2	.099		.201		.212	
F	11.639***		8.788***		8.006***	

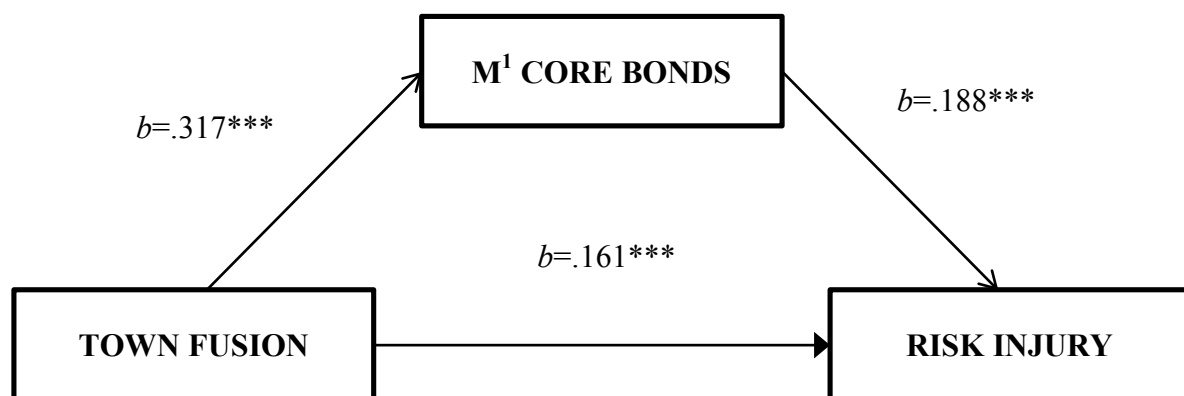
$N=215$, $*=p<.05$, $**=p<.01$, $***=p<.001$, shading added to emphasis variables with significant coefficients.

The model in the first stage significantly predicted willingness to donate time; $F(2,213)=11.583$, $p<.001$, $R^2=.099$, with the sacred value measure as the strongest contributor to the model. When the identity fusion measures were added in the second stage the model that emerged predicted a greater amount variance, $\Delta R^2=.103$, $F(4,209)=6.736$, $p<.001$, and in line with the previous analysis both fusion with local community and fusion with Japan contributed significantly to the predictive power of the model. In the final stage, *core bonds*

appeared to cancel out the contribution of fusion with local community, suggesting a possible mediating relationship (Table 4.11).

To test for this mediation, a simple mediation analysis was conducted using PROCESS (Hayes, 2012) with *core bonds* entered as a mediator between fusion with town and willingness to risk injury (Hayes, 2013: Model 4).

Figure 4-11: Core bonds mediate fusion's effect on willingness to risk injury



N=213

Direct effect $b = .101$ (.036), $p < .001$

Indirect effect M¹ $b = .060$ (.020), 95% BCa CI [.022, .100]

$\kappa^2 = .112$ (.035), BCa CI [.043, .184] [*moderate effect*]

Bootstrapping for Bias Corrected CI conducted with 1,000 Samples

The results confirmed that *core bonds* partially mediated the relationship between fusion with local community and willingness to risk injury. However, while the mediation displayed a moderate effect size, the proportion of the relationship accounted for amounted to only $b=.06$.

Collectively, these results indicated that only the two extended fusion measures were significant predictor of willingness to risk injury, when controlling for confounding factors, and that perceptions of *core bonds* mediated the relationship observed with fusion and the local community.

H4 Sacrifice Measure 3 (Voluntary Donation of Reward)

As before, the same group of variables were entered into an exploratory regression to identify potentially confounding external variables on the voluntary donation measure. This was the only true behavioural outcome in the study and thus could be argued to represent the clearest cost.

Table 4-12: Exploratory regression with voluntary donation of prize

Variable	b (SE)	Standardized β	Sig.
Constant	3129 (8580)	-	.716
Sex	-3660 (1698)	-.189	.033
Age	82.621 (58.420)	.144	.160
Residence (Yrs)	-17.481 (74.522)	-.026	.815
Town	-3800 (1867)	-.192	.044
Tengu (Pain)	1568 (868)	.221	.074
Mikoshi (Pain)	782 (939)	-.108	.407
Sacred Value 1	199 (721)	.030	.782
Sacred Value 2	1026 (679)	.156	.133
Community Imp.	19.557 (1003)	.002	.984
Total Firewalk	225 (177)	.119	.208
Total Festival	138 (92.193)	.164	.137
Affect (Mood)	214.663 (639)	.031	.738
Activation (Arousal)	--393 (351)	-.106	.265
Relative Performer	2109 (2256)	.085	.352

p<.05, p<.01**, shading added to emphasis significant predictors in the model*

Testing assumptions of the exploratory model: there was acceptable independence of residuals, as assessed by a Durbin-Watson statistic of 1.946 and no indications of multicollinearity as assessed through correlation analysis and VIF and tolerance statistics. After excluding the same 14 cases identified in the previous analysis, no additional outliers were detected. Finally, a visual examination of the resulting residuals using histograms, P-P plots and Normal Q-Q plots demonstrated a normal distribution.

Table 4-13: Hierarchical regression on costly donation

Variable	Outcome: Costly Donation					
	STAGE 1		STAGE 2		STAGE 3	
	B (SE)	β	B (SE)	β	B (SE)	β
Constant	11066 (4107)	-	3612 (4961)	-	1547 (5257)	-
Sex	-3524 (1621)	-.182*	-2654 (1630)	-.137	2899 (1598)	-.150
Town	-.4373 (1657)	-.221*	-4348 (1674)	-.220**	-4323 (1639)	-.218**
<i>Tengu</i> (Pain)	905 (591)	.128	1241 (600)	.175*	1320 (588)	.186*
<i>Tengu</i> .Fusion	-	-	922 (536)	.168	747 (529)	.136
Mikoshi.Fusion	-	-	684 (496)	.142	380 (499)	.079
Town.Fusion	-	-	-343 (442)	-.075	-707 (456)	-.154
Japan.Fusion	-	-	504 (382)	.114	563 (374)	.127
Core Bonds	-	-	-	-	1559 (608)	.245**
R^2	.086		.157		.198	
F	4.145**		3.406**		3.931***	

$N=135$, * = $p < .05$, ** = $p < .01$, *** = $p < .001$, shading added to emphasis variables with significant coefficients.

The model that emerged significantly predicted voluntary donations; $F(14,114)=2.336$, $p=.007$, $R^2=.223$, with Sex and Town emerging as significant predictors, and perceived pain of the *Tengu* performer being just outside statistical significance (Table 4.12). These three variables were thus retained as predictors for the subsequent hierarchical regression models. Based on the exploratory analysis, a three-stage hierarchical regression was conducted on donations allocated with the three items from the exploratory analysis entered as control variables in the first stage, all four identity fusion measures in the second and finally, *core bonds* added in the final stage. *Core bonds* was retained for the final stage to enable potential mediating relationships with fusion measures to be identified. All assumptions were re-examined prior to analysis but no additional outliers (values greater than 3 SD) were identified.

The model in the first stage significantly predicted willingness to donate to the shrine; $F(3,132)=4.145$, $p=.008$, $R^2=.086$, with town as the strongest contributor to the model. When the identity fusion measures were added in the second stage the model that emerged predicted a greater amount variance, $F(4,128)=2.692$, $p=.034$, $\Delta R^2=.071$, but no fusion measures contributed significantly to the predictive power of the model. In the final stage, *core bonds* emerged as the strongest predictor and still no significant predictive power was provided by the fusion measures (Table 4.13).

These results indicated that, when controlling for confounding factors, fusion was not a significant predictor of behavioural donations, and that perceptions of *core bonds* appeared to be a significant independent predictor.

Hypothesis 5 (More dysphoric experiences = Higher Fusion)

An initial spearman's correlation analysis was used to examine whether the previous experience an individual had performing as a firewalker predicted their level of identity fusion with the four fusion targets. Responses were separated into locals, (combining observers, volunteers, and performers) and non-local observers. Vendors, and low quality responses were excluded.

Table 4-14: Spearman's correlation for firewalking past experience and fusion

Sample	Variable	Tengu (Fusion)	Mikoshi (Fusion)	Town (Fusion)	Japan (Fusion)
Locals ($N=145$)	Total firewalks	.340***	.502***	.311***	.191*
Non-Locals ($N=86$)	Total firewalks	.125	.246*	.167	.035

*= $p<.05$, **= $p<.01$, *** = $p<.001$

Significant relationships with all fusion targets for local festival attendees were observed with the strongest association being the fusion measure directed at *mikoshi* performers and the weakest association with fusion with Japan. As previous participation in most cases would have involved performing as part of a *mikoshi* group such a pattern was anticipated and was also detected with a weaker magnitude in the non-local sample (Table 4.14).

Table 4-15: Hierarchical multiple regression on fusion with *mikoshi* performers

Variable	Outcome: <i>Mikoshi</i> Fusion			
	STAGE 1		STAGE 2	
	B (SE)	β	B (SE)	β
Constant	6.550 (2.199)	-	5.921 (2.126)	-
Sex	-.246 (.437)	-.056	-.027 (.427)	-.006
Age	.012 (.021)	.077	-.006 (.021)	-.040
Town (Yrs.)	.024 (.017)	.177	.022 (.017)	.166
Town	-.797 (.476)	-.170	-.775 (.457)	-.165
Tengu (pain)	-.239 (.201)	-.151	-.229 (.194)	-.145
Mikoshi (pain)	-.037 (.214)	-.023	-.080 (.206)	-.051
Sacred value 1	-.039 (.222)	-.023	-.083 (.214)	-.050
Sacred value 2	-.031 (.180)	-.021	-.023 (.173)	-.016
Comm. Import.	-.412 (.329)	-.156	-.213 (.325)	-.080
Festivals Total	-.032 (.023)	-.172	-.024 (.022)	-.130
Affect	-.273 (.164)	-.171	-.225 (.159)	-.141
Activation	-.064 (.086)	-.080	-.092 (.084)	-.116
FW Relative	-.171 (.465)	-.036	-.064 (.449)	-.013
Core Bond	.628 (.223)	.403**	.474 (.221)	.304*
Firewalk Total	-	-	.103 (.037)	.288**
R^2	.313		.372	
F	2.766**		3.319***	

$N=99$, * = $p < .05$, ** = $p < .01$, *** = $p < .001$, shading added to emphasis variables with significant coefficients.

Because a wide variety of potentially confounding factors were associated with the amount of previous firewalking performances, two exploratory hierarchical linear regressions were conducted on the local respondents' sample with the first using fusion with *mikoshi* as the outcome and the second fusion with local community. Factors entered

as control variables included the demographic factors- age, sex, residence time, town; mood; activation; times attended festival; two sacred values measures; importance of festival to community; perceptions of *core bonds*; and presence of family or friend in firewalkers. Testing assumptions there was acceptable independence of residuals. There was also no indication of multicollinearity in either model, as assessed through correlation analysis and VIF and tolerance statistics. Furthermore, an examination of standardized residuals and leverage measures identified no extreme outliers or cases with undue influence. Finally, a visual examination of the resulting residuals in histograms, P-P plots, and Normal Q-Q plots demonstrated normal distributions.

Table 4-16: Hierarchical multiple regression on fusion with local community

Variable	Outcome: Community Fusion			
	STAGE 1		STAGE 2	
	B (SE)	β	B (SE)	β
Constant	6.768 (2.103)	-	6.482 (2.106)	-
Sex	-1.110 (.418)	-.266**	-1.010 (.423)	-.242*
Age	-.059 (.020)	-.409**	-.067 (.021)	-.465**
Town (Yrs.)	.020 (.017)	.157	.019 (.016)	.152
Town	-.182 (.455)	-.041	-.172 (.453)	-.038
Tengu (pain)	-.362 (.192)	-.240	-.358 (.192)	-.237
Mikoshi (pain)	.159 (.204)	.105	.139 (.204)	.092
Sacred value 1	-.059 (.212)	.037	.039 (.212)	.024
Sacred value 2	.117 (.172)	.084	.121 (.171)	.087
Comm. Import.	-.102 (.315)	-.040	-.012 (.322)	-.005
Festivals Total	.025 (.022)	.143	.029 (.022)	.163
Affect	.107 (.157)	.070	.129 (.157)	.084
Activation	-.014 (.083)	-.018	-.027 (.083)	-.036
FW Relative	-.664 (.445)	-.146	-.615 (.445)	-.136
Core Bond	.297 (.213)	.200	.227 (.219)	.153
Firewalk Total	-	-	.047 (.036)	.137
R^2	.314		.327	
F	2.776**		2.723***	

$N=99$, * = $p < .05$, ** = $p < .01$, *** = $p < .001$, shading added to emphasis variables with significant coefficients.

The results from these two paired exploratory regression analyses, indicated that previous firewalking experience was a significant predictor for local fusion targeted at other firewalking performers, when controlling for a host of potential confounding variables. Conversely, previous performances did not significantly predict fusion with the local community suggesting that the bonding effect of shared dysphoric experiences had a limited relational range.

Hypothesis 6 (More pain = Prosocial behaviour from ingroup observers)

To examine if perceptions of pain were associated with increased prosocial behaviour for observers, a spearman's correlation analysis was conducted between the two perceived pain measures, the three prosocial measures, and the single social identification item. Responses from local and non-local observers were analysed separately as the hypothesis predicted effects to be observed with ingroup members. Responses for performers were also included for exploratory purposes.

Table 4-17: Spearman's correlation matrix for pain and prosocial measures

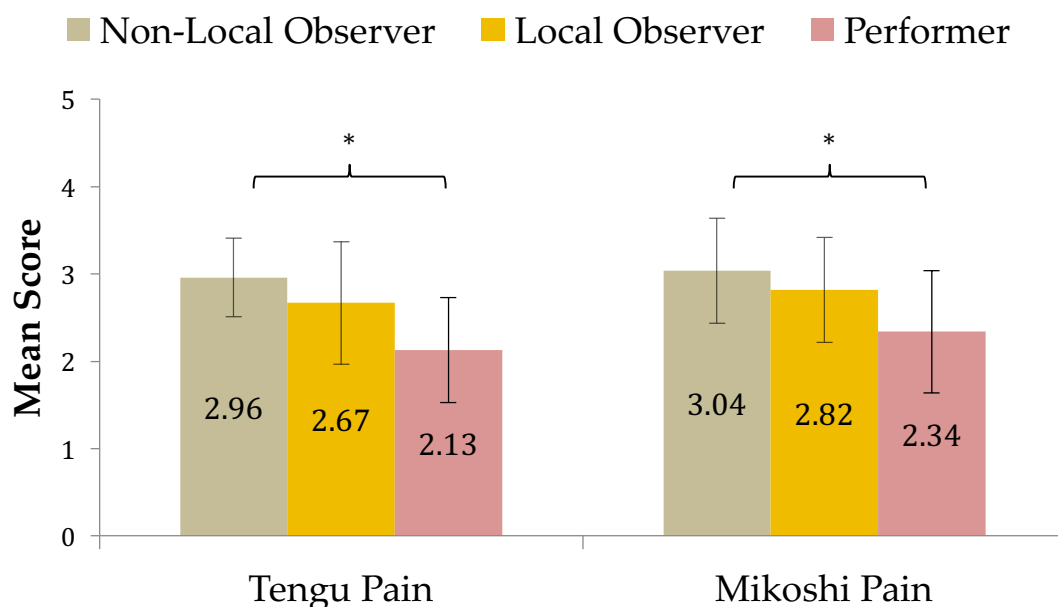
Sample	Variable	Donate Time	Risk Injury	Prize Donation	Identit.
Local Observer (N=67)	Tengu (Pain)	-.049	.172	.204	.195
	Mikoshi (Pain)	-.028	.088	-.010	.093
Non-Local Observer (N=81)	Tengu (Pain)	-.235*	-.052	.228	-.131
	Mikoshi (Pain)	.179	-.118	.039	-.052
Performer (N=35)	Tengu (Pain)	-.142	-.091	-.052	-.406*
	Mikoshi (Pain)	-.206	-.093	-.099	-.400*

*= $p < .05$, **= $p < .01$, *** = $p < .001$

The results indicated only a single significant negative correlation for non-local observers between the perceived pain of the *Tengu's* performance and reported willingness

to donate time to help the shrine. But given the number of comparisons conducted and the level of significance ($p=.043$) it is likely that this result is due to random variation. For performers, a repeated negative correlation with perceived pain and local identity identification was found. This may indicate that performers with little experience or weaker connections to the local community anticipated the event to be more painful. This interpretation is supported by performers rating both the *Tengu*, $F(2,189)=9.602$, $p=.005$, and the *mikoshi* performers, $F(2,182)=3.258$, $p=.041$, fire walking experiences to be significantly less painful than the observers rated the same experiences (Fig.4.12).

Figure 4-12: Perceived pain by observers and performers



**= $p<.05$, based on Hochberg GT2 post-hoc comparisons.*

Note: Error bars represent standard deviations.

In summary, counter to predictions greater perceptions of pain were not associated with increased prosocial behaviour or more inclusive social identification for ingroup and outgroup observers and performers.

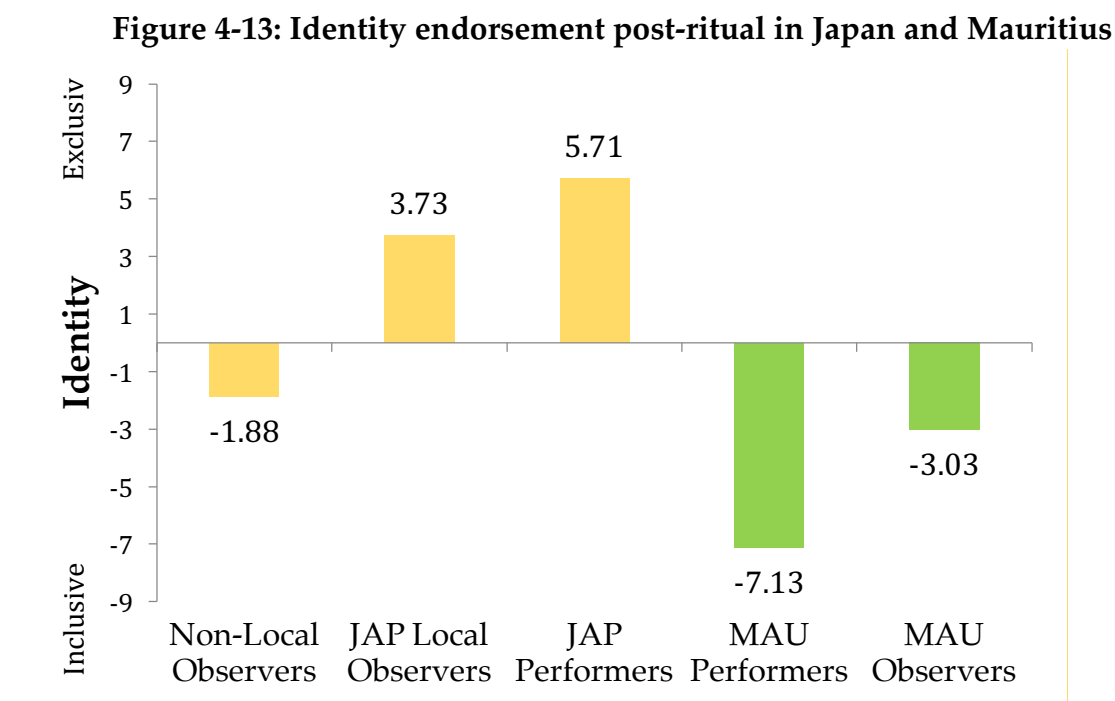
4.5 Discussion

This study represents the first independent replication to address the findings reported in Xygalatas et al. (2013b). Two key findings were examined concerning the effects of observing and participating in extreme ritual events; the first, that prosocial behaviour directed at the ingroup is increased and the second, that more inclusive social identities are promoted.

In regards to prosociality (H1), the results obtained overall offered no strong support for an increase. On the behavioural donation measure, which was the closest to the measure employed by Xygalatas et al. (2013), no significant difference in donations was observed between the pre- and post-periods for either observers or performers. Observers and performers also demonstrated no greater willingness to donate their time to help the local shrine after observing the firewalking performance. However, for self-reported willingness to risk injury to assist a firewalker who was in danger a marginally significant increase was detected amongst firewalking performers after they had just performed their firewalk. Responses from observers displayed a similar significant increase in willingness to risk injury to intervene but this relied on pooled data that did not distinguish between local and non-local observers and when examined as independent groups there was no statistical difference in responses between periods. This suggests that the finding is more likely to be an artefact from there being a larger proportion of local observers with stronger relational and familial ties with firewalkers in the post event period (local observers- 84.7%, $N=83$ vs. non-local observers- 13.9%, $N=15$). Unfortunately, this was an unavoidable sampling bias as

non-local observers left the festival area almost immediately after the firewalking performances were completed.

Xygalatas et al. (2013b) also operationalised endorsement of a more inclusive national identity over an ethnic religious identity as an indicator of prosociality. This has validity in the context of Mauritius as a previous study linked endorsement of national identity over ethnic religious identity as being associated with more positive views towards outgroups (Tseung-Wong & Verkuyten 2010). In the context of Japan, which has very low levels of ethnic diversity and syncretic religious traditions, the same contrast was inappropriate. Therefore, a comparable measure that contrasted local (exclusive) and national (inclusive) identities, was employed. Using this measure, with the same interval scale as Xygalatas et al. (2013b), a contrasting pattern was found (Fig.4.13).



Note: Japanese data is based on post-firewalk period; Mauritius data is from High Ordeal ritual condition (Xygalatas et al. 2013b)

As Fig 4.13 indicates, local performers and observers in the Japanese sample more readily endorsed the more exclusive local identity. Moreover, the same pattern with lower exclusive identification was also evident in the pre-firewalk period (H2). It was only non-local observers who displayed a result similar to that found in the Mauritius sample. This suggests that the finding from Mauritius may be context specific which Xygalatas et al. did recognise as a possibility (2013b: 3). It is also important to recognise that the Thaipusam festival in Mauritius is a *national* festival even though it is primarily associated with the Hindu Tamil ethnic community. The *Tengu* firewalking festivals in Shakotan and Furubira in contrast are clearly local events that are promoted as part of a unique regional identity. This can be seen in the tourist material produced for these towns, and in various physical representations, such as a large mural of a *Tengu* painted on the side of a building in Furubira.

In this context, it does not seem surprising that observing or taking part in the firewalking festival would promote a sense of local self-identification. But the important point here is that the results suggest it is clearly too early to conclude that collective dysphoric rituals typically promote an inclusive form of prosociality. It seems instead more likely that whether responses are more parochial or inclusive will be primarily determined by the salient norms associated with particular groups. In the case of Mauritius, it could also be argued that since Hindus represent the majority ethnic group endorsing a national identity during a national Hindu festival and choosing to donate money to a Hindu temple do not necessarily represent inclusive prosociality. Whatever the case, the present findings

provide an important independent sample that did not offer any strong support for extending claims about extreme rituals and prosociality more generally.

The third hypothesis addressed recent predictions from Modes theory that the intense relational bonds of identity fusion are generated, or at least promoted, through collective dysphoric rituals. This prediction was examined using the visual identity fusion measure with a variety of different target groups. The visual measure is generally considered less accurate than the verbal measure however, in the field the visual item was much quicker to administer and proved significantly less difficult for respondents to understand. The Japanese translations for the identity fusion scale produced for the experiments in Chapter 3 were found during pre-testing to be difficult for local respondents to answer as Japanese concepts of self-identity do not map entirely coherently onto the expressions used in the scale items. Although this does raise concern for the cross-cultural validity of the measure it should be noted that the scale had already been translated for use with a Japanese sample (Swann et al., 2014) and displayed high levels of internal reliability in the experimental study conducted with students reported in Chapter 3. Additionally, the wording of the fusion scale even in English is somewhat unusual including statements such as 'My group is me', so the issues identified might apply if attempting to administer the English version of the scale in rural communities with older populations. The decision to exclude the verbal scale also does not impact comparisons with Xygalatas et al. (2013b) as this measure was not included in the original study.

Returning to the results collected from the visual measure, comparisons conducted between response groups in the post-event period revealed that, in accordance with

expectations, firewalking performers had significantly higher levels of fusion with other performers than observers. Furthermore, examining the *core bonds* measure, performers were also found to have significantly higher scores than both categories of observers. Tellingly, no differences were detected in overall levels of fusion with the local community or Japan. This finding increases confidence that the differences detected are valid, rather than an artefact of a response set such as performers displaying a general tendency to respond more strongly to scale items.

Overall, when comparing response groups, the evidence follows the patterns expected if collective *performance* is crucial to local identity fusion. That is, compared with mere observers, firewalking performers displayed significantly higher levels of fusion with all other firewalking performers (*Tengu* & *Mikoshi*) and a stronger sense of relational connections and personal consequentiality (*core bonds*). Fusion with the extended community and national identity was also elevated but for these measures, the elevation was matched by observers.

Alternatively, this result could be due to selection bias with firewalking performers being the most committed strongly committed group members overall. However, this interpretation is called into question by the fact that approximately 20% of local observers were also previously firewalkers yet significant differences were detected between this subgroup and performers in terms of fusion with *mikoshi* performers ($p < .001$) and the *Tengu* performer ($p = .017$). These results however do not completely rule out firewalking performers representing a particularly devoted type of local community member. And one piece of evidence which supports this is that in the observers' sample those with previous

firewalking experience did not differ in the distribution of the *core bonds* responses ($p=.196$), whereas those without experience displayed a significantly lower distribution ($p<.001$).

Examining the changes in fusion between periods helps to clarify distinctions between the different categories of respondents as non-local observers displayed significant increases in fusion with the *Tengu* performer and the local community, local observers displayed a decrease in levels of fusion with both *Tengu* and *mikoshi* performers but an increase in levels of fusion with local community and country, and performers displayed no significant change on any measure but were already approaching ceiling levels for fusion with every fusion target group in the pre-event period excluding the *Tengu* performer. The most intriguing finding from these results is the drop in fusion with performers detected for local-observers and the converse increase in fusion with categorical affiliation groups (community and nation).

This suggests that for local observers watching the dysphoric performances increases their perception of distance from the firewalkers, but conversely motivates them to reaffirm their ties with the local community and the Japanese nation. If this result is valid it implies that dysphoric performances may serve two contrasting purposes which could serve to synergistically increase cohesion amongst a given community. That is, performers of collective dysphoric rituals appear to develop a unique, visceral perception of connection with other firewalkers who share the same experience, which this is accompanied by strong relational bonds to the relevant wider community identity. Ingroup observers in contrast display significantly lower levels of connection with firewalking performers (at least

temporarily) but compensate for this by affirming their relational connection with their various social identities.

This account accords with the emphasis in Modes theory on the intense bonds forged amongst performers in imagistic rituals (Whitehouse & Lanman, 2014) but also supports a costly signalling interpretation (Bulbulia & Sosis, 2011; Soler, 2012) that dysphoric performances may generate compensatory affiliative responses being made by ingroup members who are not sacrificing for their group. Although it would be tempting to use this finding to weave a narrative about the importance of such a system in forging more cohesive communities throughout evolutionary history, it is important to note that no significant differences were detected between groups in the post-event period on any of the prosocial outcome measures. That is, if it is true that local observers experienced a higher level of identity fusion with their local community identity after observing ingroup members complete a dysphoric ritual, this did not translate into any significantly greater prosocial behaviour, even when compared against outgroup members⁶⁶.

Another alternative is that the individuals taking part in firewalking events, predominately rowdy local youths, are seen by observers, following their chaotic and dangerous performances, as being characteristically different from them. However, this interpretation seems difficult to defend given the almost universal positive regard expressed for the firewalkers during conversations with locals in Shakotan and Furubira

⁶⁶ To test this additional Kruskal Wallis tests were conducted between response groups for all three prosocial outcomes but no significant differences were detected for any of the measures.

and furthermore fully 70% of the local observers had a friend or family present amongst the firewalkers.

For non-local observers, fusion with *Tengu* performer and local community increased, however this may be an artefact created by bias in the only sample of non-locals who were available to complete questionnaires in the post-event period being those with some additional ties to the community in question. This interpretation is supported by the proportion of non-local observers responding that they had a friend participating in the firewalking performance this year, which increased from 21.1% in the pre-event period to 41.7% in the post-event period. The proportion that indicated they had a family member participating also increased from 2.8% to 8.3%. This interpretation is partially supported by comparing the fusion results of non-local observers and local observers in the post-event period, as no significant difference was detected in fusion levels reported with the mikoshi performers; $t(44)=.702$, $p=.487$ and the local community; $t(44)=-.957$, $p=.344$. However, there was a statistically significant difference in fusion level with the *Tengu*, $t(44)=2.659$, $p=.011$, $d=.802$ and furthermore the small sample size of non-local observers in the post-event period ($N=15$) reduces confidence in these findings.

To sum up, the results provide tantalising preliminary evidence for differential responses in group bonding targets between performers and observers. The predictions of Modes theory about the importance of visceral shared experiences seem to be supported and were also reinforced in the narrative accounts provided by performers throughout the town during informal interviews. Many previous performers who were not taking part in the current year's performances were quick to identify permanent burn scars that they had

obtained during previous performances and to recount their experiences vividly. There were also those who expressed some doubts about whether their injury had been worth it, or was merely the result of misplaced youthful exuberance. There is no way to disentangle such responses from the widespread tendency for Japanese people to self-depreciate (Akimoto & Sanbonmatsu, 1999; Heine, Lehman, Markus, & Kitayama, 1999), especially when talking with outsiders, but it could perhaps also signal a progressive weakening of the bonds created through shared dysphoric experiences. Further evidence is required from different dysphoric ritual communities before any broader conclusions can be drawn.

Another unexpected finding (H4) from the current study was that extended fusion with country was found to be a predictor for two of the three local orientated prosocial outcomes. For the willingness to donate time measure, it was the only significant fusion target that proved predictive and for willingness to risk injury it was only slightly less predictive than fusion with local community. Furthermore, these contributions appeared to be channelled through some other pathway than relational bonds, as when the *core bonds* measure was entered into the regression models the coefficients observed remained unaffected. In contrast, the connection between fusion with local community and willingness to risk injury was found, as predicted, to be mediated by perceptions of *core bonds*. This finding replicates the results reported in Chapter 2 but its occurrence in only one outcome measure makes it at best a mixed result. For the behavioural measure the results also revealed that fusion was not a significant predictor of this outcome when controlling for confounding factors.

Collectively, these results offer only limited support for H4 as counter to expectations it seems that extended fusion measures were the most significant predictors of prosocial outcomes. The willingness to risk injury outcome, which is potentially a more valid operationalization of a costly sacrifice, did indicate that levels of community fusion were significantly predictive of the outcome. However, it is the complete absence of the local performer fusion targets from predictive models of the prosocial outcomes that is most surprising, especially when the injury scenario involves explicit harm to a firewalker. One explanation could be that the relationship is diluted by the post-event decrease in fusion with local performers by local observers reported in H3. However, post-hoc regressions which excluded local observers still did not identify the anticipated contribution from local fusion targets. The validity of these findings require further exploration in independent samples.

Another piece of evidence that offers support to the Modes theory account of local fusion processes was that the total amount of previous firewalking experiences was a significant predictor for local fusion targeted at other firewalking performers (H5) when controlling for a host of potential confounding variables. This is not a surprising result but is interesting to note that previous performances did not significantly predict fusion with the local community, suggesting that the bonding effect of shared dysphoric experiences had a limited relational range.

Finally, focusing on the effect of perceptions of pain in firewalking performances (H6), no correlations were found with any response group on prosocial measures or inclusive social identification. This contrasts with the findings of Xygalatas et al. (2013b: 3) that

higher levels of perceived pain were “directly associated with greater donations”. Also, counter to the findings in Mauritius, perceptions of pain differed between performers and observers, with performers reporting experiences to be significantly less painful. The discrepancy may reflect a difference in the contextual motivation for the endurance of pain. In particular, in Thaipusam enduring painful piercings is associated with imploring the Hindu God of War (*Murugan*) for assistance and thus has clear devotional connotations. For the firewalking festivals performed in Shakotan and Furubira however there are almost no explicit doctrinal motivations and individuals participating are not regarded as being devotees to any Shinto deity. The focus as with most festivals and religion in general in Japan is on the spectacle and the local community. Whether this difference in connotation can account for the null finding or whether there is another cross-cultural explanation will again require further cross-cultural data to determine.

Conclusion

This study represented the first independent field replication of the effects on performers and observers of extreme rituals reported in Xygalatas et al. (2013b). The results obtained from the Japanese sample however, provided only limited support for the prediction that prosociality towards the ingroup would be higher after an extreme ritual performance. Specifically, a marginal trend was found amongst firewalking performers to report a greater willingness to risk injury to help another firewalker in the post-event period but this effect was not replicated amongst observers, when divided into locals and non-locals. Visual inspection of the voluntary donation outcome also suggested the anticipated boost

in the post-event period but this was not confirmed by statistical analysis due to wide standard deviation in responses.

It is important to recognise that the measures used in both studies differed slightly due to contextual factors and therefore this study does not represent an exact replication. However, the current study was intentionally designed to replicate Xygalatas et al.'s (2013b) methods closely, including by providing respondents the chance to anonymously donate any amount of a large reward sum to the local shrine. That the money donated in the current study was hypothetical, as it relied on the respondent being selected as a random winner, represents a significant change. But it seems likely that if this was a factor it would have served to increase the generosity baseline generally across the sample. Furthermore, all indications in the field suggested that respondents did treat this measure seriously and this was supported by their willingness to provide personal contact details, including mobile telephone numbers, in case they were selected. Whatever the case for the donation measure, the prosocial self-report measures also failed to detect any substantial increase in responses from performers and observers after the ritual performances.

Second, on the social identification measure a contrary pattern to the Mauritius data emerged of heightened endorsement of local and exclusive identities over more inclusive national identification in the post-event period. This finding was discussed in the preceding section and there are valid reasons to anticipate a different pattern in Japan. But the finding still reinforces the view that it is too early to assume that extreme rituals will necessarily promote inclusive identification processes.

In addition to partially replicating Xygalatas et al. (2013b), this study was also intended as the first examination in a religious ritual context of the recently revised predictions by Modes theory concerning bonding in high arousal dysphoric ritual events and the central role of identity fusion in such processes (Swann et al., 2012; Whitehouse & Lanman, 2014; Whitehouse et al., 2014). Broadly, the results obtained were supportive of the theory, with firewalking performers displaying significantly higher levels of *local* fusion with co-performers and an elevated perception of *core bonds* (H3).

More previous experience as a firewalker performer also predicted levels of fusion with other relational fusion targets (H5). Conversely, local performers demonstrated an increase in *extended* fusion with the local community and broader national identities in the post-event period. This result may also be partly accounted for by the event making perceptions of differential roles salient but even if this were the case that could be part of the underlying mechanisms generating differential bonding patterns.

These results provide preliminary evidence for differential responses in group bonding targets between performers and observers and imply that theories addressing dysphoric collective rituals need to take greater care in distinguishing predictions on the affiliative tendencies of observers and performers of imagistic rituals. The results also accord with costly signalling theories that posit collective dysphoric ritual performances as a 'social technology' (Fischer & Xygalatas, 2014) that can motivate observers to increase their affiliation and display stronger devotion to a relevant group identity (Bulbulia & Sosis, 2011; Henrich, 2009; Lanman, 2012).

The current study reaffirmed the well-established connection between identity fusion and costly sacrifice. But, surprisingly, it was the *extended* fusion target of national identity which proved to be the most consistent fusion measure, predicting significantly two of the three prosocial outcomes. Fusion with local community was a more significant predictor of willingness to risk injury, the most visceral of the costly sacrifice measures employed and this connection was found to be partially mediated by perceptions of *core bonds*, in accordance with the findings reported from the Brazilian Jiu Jitsu data in Chapter 2. However, the original total effect between community fusion and willingness to risk injury was relatively weak ($b=.161$, $p<.001$) and, as such, this result should not be over-interpreted, particularly when no comparable pathways were found with the willingness to donate time and behavioural donation measures.

Finally, predictions derived from costly signalling accounts (Bulbulia & Sosis, 2011; Sosis & Alcorta, 2003) and recent field (Xygalatas et al., 2013b) and experimental studies (Bastian, Jetten, & Ferris, 2014) that higher perceptions of pain should be associated with subsequent stronger prosocial attitudes and behaviour were not supported by the current data.

For the current thesis this chapter offers some much needed ecological validity to the data collected about dysphoric rituals and group bonding. The dysphoric ritual performances were not associated with religious devotion per se but the significant associations observed with the *core bonds* measure demonstrate that these events were personally consequential for many performers and observers. The critical issue identified in previous chapters of subjective positive assessment of dysphoric experiences was not

addressed explicitly in this chapter but in interviews with present and previous performers it was clear that the vast majority regarded the events as euphoric experiences, with pain or potential injury being dismissed as an insignificance.

It is also worth noting that, at least for the male performers, a sizable number are intoxicated and this likely reduces fear and decreases sensitivity to pain. As an illustration of this, the current *Tengu* performer in Furubira, who does not drink during or before his performances, was accorded special admiration with one shrine authority and a previous *Tengu* performer laughing that they could not imagine enduring the fire without being drunk. As with the Brazilian Jiu Jitsu promotion experiences, subjective positive assessments do not alter the physiological reality of the dysphoric experience but it seems that post event euphoria should be given more consideration as a possible motivating factor for the bonding and prosocial outcomes that are predicted following dysphoric ritual performances.

Limitations

Although conducting research in the field can provide unique data that has much stronger ecological validity than results obtained from controlled experiments, it also carries with it a number of unavoidable limitations. The most significant of these is an inability to control for various random confounding factors, including things such as the surrounding environment factors in which people responded to questionnaires. Collecting questionnaires from the gathered audience in the pre-firewalking period meant that a significant portion of responses were obtained from short term visitors and photographers, most of who had little connection with the town. Although this represents important data

since it can provide a baseline for local responses, it also proved impossible to collect a similar post-event sample. Most of the short term visitors had arrived by car or on a tour bus and they were thus eager to leave as soon as the festival had concluded. This pattern was observed during the first night of the festival in Shakotan and subsequently efforts were made to try to avoid collecting too many responses from short term visitors and to request, in advance, for those who were on a more flexible schedule to agree to complete questionnaires in the post-event period. Even with such efforts, only 15 usable responses from non-locals were collected in total in the post-event period in both towns.

Another factor that limits how the relationships observed can specifically be attributed to processes of identity fusion is that no measure for categorical group identification was employed. This decision was taken due to a desire to avoid verbal scales and keep materials simple but it meant that it is not clear that the effects attributed to identity fusion could not be also accounted for by categorical group identification. This is a critical point given the unexpected role uncovered for categorical group identification in forming relational bonds in Study 2 in Chapter 3. Whether the bonds characteristic of fusion are really the most relevant bonding mechanism for performers or observers of extreme rituals can only be properly addressed by employing measures that directly contrast categorical and relational ties. The difficulty here will be in selecting measures that are simple enough to be employed in the field but precise enough to distinguish between the two related constructs of group identification and identity fusion.

Finally, an argument could be made that the firewalking festival as an entire event represents an extreme ritual, and that therefore it is invalid to use the pre-performance

period as being analogous to Xygalatas et al. (2013b)'s low arousal comparison group. There is some validity to this objection but it should be noted that the low arousal ritual context in the Mauritius study also represented one part of *Thaipusam* celebrations. Furthermore, fine-grained studies of physiological arousal amongst performers and observers, also collected by Xygalatas and colleagues (Fischer et al., 2014; Xygalatas et al., 2011) at other firewalking festivals demonstrate clearly that the act of observing and participating in the actual performance represent a dramatic and affecting experience. It thus seems unlikely that the pre-performance time period is not significantly different in character from the post-performance period.

CHAPTER 5: CONCLUSIONS

General Conclusions

Through the preceding chapters this thesis has provided a raft of novel evidence concerning the individual psychological impact and socially orientated consequences of collective dysphoric rituals. Data was explored from martial artists (Chapter 2), participants in two artificial ritual experiments conducted in the UK and Japan (Chapter 3), and performers and observers of Shinto firewalking festivals in Japan (Chapter 4). In this chapter the data collected is synthesised to address the four core research questions identified at the beginning of the thesis.

Q1. Do collective dysphoric rituals promote affiliative bonds, and if so, do effects differ amongst participants and observers?

Recent formulations of modes theory have posited a key connection between processes of identity fusion and shared dysphoric ritual experiences (Swann et al., 2012; Whitehouse & Lanman, 2014). In this thesis for the first time this connection was examined directly across a variety of ritual contexts, including actual religious dysphoric ritual performance. Taking the results from the martial arts data in Chapter 2 first, hypothesised connections between dysphoria and identity fusion (H1a, H2) and the formation of relational bonds (H1b, H5b) were generally not supported. Specifically, relational bonds and levels of fusion were not found to be significantly greater amongst individuals who experienced dysphoric belt whipping gauntlets during promotion events on both verbal and pictorial measures of fusion (H1). Furthermore, examining negative promotion experiences across the full sample of BJJ practitioners, no significant correlation was observed with identity fusion and a

hierarchical regression found that it negatively interacted with the presence of belt whipping to predict lower fusion scores (H2). Negative promotion experiences and relational bonds also demonstrated no robust relationship (H5b). Furthermore, in contrast to the lack of associations between dysphoric experiences and fusion, positive experiences displayed a robust relationship with identity fusion, even when controlling for group identification. Fusion was also found to mediate the relationship between positive experiences and costly sacrifice outcomes.

In Chapter 3 both of the experimental studies found no evidence of elevated fusion in the dysphoric condition. This outcome however may be a result of methodological limitations, specifically it may be unrealistic to expect groups of naïve participants performing an unfamiliar ritual in a laboratory setting with a group of strangers to be capable of generating the intense relational bonds characteristic of identity fusion. Modes theory suggests that personally consequential, autobiographical memories formed from shared dysphoria are the key ingredient in establishing identity fusion amongst ritual members but in the experimental setting personal consequentiality was clearly lacking in the artificial rituals performed.

In contrast to the martial arts and experimental studies data, the results from the firewalking field study provided evidence for a connection between fusion and dysphoric rituals (H3). Specifically, it was found that for performers participation increased perceptions of *local* fusion with other performers, while conversely for local observers' fusion was elevated with *extended* fusion targets. This represents the first evidence for

differential affiliative consequences of dysphoric ritual performances amongst individuals all of whom held local commitments but performed different roles.

However, it should also be noted that performers' fusion with the *extended* community target was already high in the comparative baseline pre-event period. This makes it unlikely that any significant positive shift would be detected. Despite this limitation, the results did still demonstrate a differential affiliative response between performers and observers and this was argued to provide support for recent predictions of Modes theory regarding the bounded nature of the relational bonds forged in shared dysphoric ritual experiences.

Alternatively, these results could also be interpreted as offering evidence of a 'costly signalling' (Bulbulia, 2004a; Henrich, 2009; Irons, 2001) mechanism in which ritual observers are encouraged to increase their commitment and sense of connection with a relevant group identity or community due to the CREDs of other group members. This interpretation is supported by the drop in fusion with performer targets by local observers, which could indicate guilt or unconscious resentment generated by the performers' public display of their devotion to the community. However, such an interpretation remains highly speculative in light of the limitations of the sample and visual measures collected. Furthermore, the firewalking data does not offer the opportunity to directly examine the significance of categorical bonds as no comparable measure of group identification was collected as a result of the limits of field data collection. Therefore, the findings reported in Chapter 4 represent a preliminary finding that is in need of verification through

independent replications at different field sites, ideally using more fine grained measures of fusion and group identification.

Collectively, the majority of studies failed to find the anticipated connections between dysphoria, fusion, and relational bonds. However, as the positive findings were derived from the most ecologically valid context, which represents the only study in the thesis that is associated with a religious tradition and evocative cultural symbolism, the negative results are significantly qualified and in need of further investigation. As has been previously argued by Sosis and Bressler (2003), it may be that the ability of religion to infuse events with a sense of sanctity and tie them to supernatural ideologies makes dysphoric rituals performed in religious contexts especially effective at generating rich symbolic experiences to fuel the introspective reflection regarded as a key process for building enduring relational bonds in imagistic rituals in Modes theory. Rossano (2012) has emphasised the ability of rituals to serve as vehicles to transmit and emotionally tie individuals into group norms and values and this also might account for why dysphoric collective rituals seem to be more impactful when associated with elaborate religious systems.

Table 5-1: Summary of Findings for Core Q1

Do collective dysphoric rituals promote affiliative bonds, and if so, do effects differ amongst participants and observers?				
	BJJ Promotions	Artificial Ritual Study 1	Artificial Ritual Study 2	Shinto Firewalking
Overall	Negative	Negative	Negative	Positive
Summary of Findings	No evidence of increased fusion or identification with more dysphoric experiences.	No evidence of increase in relational bonds in the dysphoric condition.	No evidence of increase in relational bonds in the dysphoric condition.	Higher levels of fusion observed in the post event period for performers and observers. Targets for the increase varied between performers and observers.

Q2. Does participation in, or observation of, dysphoric rituals generate prosocial behaviour directed towards the ingroup?

The strongest current evidence for the prosocial effect of dysphoric rituals was provided recently by a study conducted by Xygalatas et al. (2013b) with Hindu piercing rituals in Mauritius. The study reported in Chapter 4 focusing on Shinto firewalking festivals represented the first attempted independent replication of these findings. Focusing on the measures which were most similar to those used in the original study; there was no evidence of a significant increase in voluntary prosocial donations directed towards the local shrine in the post event period. Furthermore, accepting the operationalization of inclusive social identification as a measure of prosocial attitude from Xygalatas et al. (2013b), a contrary pattern was observed to that previously found in Mauritius. Specifically, performers and local observers endorsed more *exclusive* social identities following the firewalking event and only non-local observers displayed a greater endorsement of an inclusive national identity.

Two novel self-report measures that represented a willingness to endure a cost (donating time or risk an injury) to benefit the ingroup were also assessed. For willingness to donate time, no change was observed for performers or observers in the post event period, but for willingness to risk injury there was a significant increase with observers and a marginal positive trend for performers. Yet, as discussed in Chapter 4 this effect, at least for observers, is likely to have been the result of sampling bias. Consequently, the evidence from the firewalking field study failed to lend support for previous findings of heightened prosociality following dysphoric ritual performances.

Turning to evidence from the other chapters, null results were also returned from the martial arts data as across three independent costly sacrifice measures in that no association was found with either experience of belt whipping gauntlets or more broadly with previous negative experiences in promotion practices (H5a). The picture that emerged from the experimental studies was more mixed: Study 1 found a marginal increase in cooperative behaviour amongst participants in the dysphoric condition (H2), and Study 2 although finding no ingroup preference in intergroup allocation tasks did demonstrate increased preference for fictional group members in a BIAT (H2). Significantly, those in the euphoric condition alternatively demonstrated a stronger preference for other group members in a forced ranking question.

Overall then, the findings are somewhat inconsistent, results from the more ecologically valid field site study and martial arts groups were consistently negative, whereas the experimental studies provided mixed results. Further studies are clearly required to establish whether a broad association exists or whether prosociality is tied to specific group norms and particular cultural contexts.

In relation to questions specifically about the role of pain in promoting prosociality (Bastian, Jetten, Hornsey, et al., 2014; Fischer & Xygalatas, 2014; Xygalatas et al., 2013b) the evidence from the current studies were more unequivocally negative. In both the Shinto firewalking (H6) and the martial arts data (H5a) there were no correlations observed between any of the prosocial outcomes and perceptions of performers' pain.

Table 5-2: Summary of Findings for Core Q2

Q2: Does participation in, or observation of, dysphoric rituals generate prosocial behaviour directed towards the ingroup?				
	BJJ Promotions	Artificial Ritual Study 1	Artificial Ritual Study 2	Shinto Firewalking
Overall	Negative	Positive	Mixed	Negative
Summary of Findings	No correlation with dysphoric experiences and three sacrifice measures	Marginal increase in cooperation in dysphoric condition	- No increase in ingroup donations in dysphoric condition. - Elevated implicit positive associations with fictional group members.	- No increase in voluntary donations. - More exclusive identity selected

Q3. How do collective dysphoric rituals identities impact self-identity?

This question ended up being partly addressed in the summary provided for Q2, as the self-identification measure has also previously been interpreted as a measure of prosociality. The relevant finding from the firewalking fieldwork studies suggested that they promote exclusive, local identities but it was argued that such processes will be highly contextual and potentially culturally specific. It would be misleading to conclude on the evidence provided in this thesis that exclusive social identities are necessarily promoted by shared dysphoric experiences. Again further research is required to explore the potential role of religious devotion and group norms invoked by salient identities in ritual contexts. Japan represents an unusually homogenous cultural environment and thus it may be that participation in unique regional festivals help to promote a sense of distinctiveness (Brewer, 1991; Leonardelli, Pickett, & Brewer, 2010).

The Tajfel allocation matrices employed in the Study 2 of Chapter 3 were intended to help examine variations in ingroup preference and test for motivations in intergroup discrimination. However, no significant differences were found between conditions. This,

coupled with previous experimental evidence from Japan in minimal group settings (Jin et al., 1996; Karp et al., 1993; Yamagishi, Mifune, Liu, & Pauling, 2008) and the findings of elevated implicit ingroup preferences in the dysphoric condition and relational ingroup preference in the euphoric condition seem to support the position that ingroup preference and intergroup discrimination are not hydraulic processes. However, further investigation is required before any strong conclusions can be offered on this question.

Table 5-3: Summary of Findings for Core Q3

Q3: How do collective dysphoric rituals identities impact self-identity?				
	BJJ Promotions	Artificial Ritual Study 1	Artificial Ritual Study 2	Shinto Firewalking
Overall	--	--	Mixed evidence for ingroup preference	Exclusive
Summary of Findings	--	--	Some evidence for enhanced ingroup preference invoked by shared dysphoria but accompanied by an absence of intergroup discrimination.	Identification with exclusive identities higher for local observers and performers. Inclusive identities found with non-local observers.

Q4. Are the individual and social effects generated from collective dysphoric rituals distinct from those of collective, high-arousal euphoric rituals?

Only the second experimental study explicitly contrasted a high arousal euphoric condition with a high arousal dysphoric condition and on most outcomes no clear significant difference was detected. However, on subtler measures of ingroup preference (the BIAT and resource allocation target) a significant difference was observed between the two conditions that suggested respectively a heightened positive association with an allocated group identity in the dysphoric condition and a greater preference for relational ingroup members in the euphoric condition. These findings are preliminary and offer

mixed support for the proposal of a distinctive role in shared dysphoric arousal in generating group affiliation (Whitehouse & Lanman, 2014) but they lend credence to the argument that high arousal imagistic rituals as a collective category are important for developing group bonds. The effects observed cannot also be readily observed by existing theories, such as cognitive dissonance, as group allocations were made arbitrarily. Furthermore, the potential confound of reciprocal self-interest (Yamagishi et al., 2008) was controlled for in the study design (Yuki & Yokota, 2009).

Turning to more general processes of *subjective* positive assessment of experiences, the BJJ study found unexpected but highly significant results for a clear relationship between *positive* experiences and identity fusion. The results offered initial evidence that identity fusion could function as a core mediating mechanism in translating positive ritualised experiences, and potentially euphoric arousal, into pro-group outcomes and stronger relational bonds. A relationship that is currently not emphasised in the identity fusion literature (Swann et al., 2014, 2009, 2012). Following this finding the artificial experimental studies explored these associations and collectively cast some doubt over the generalisability of these results. In particular, in Study 1 the expected association with higher levels of fusion was not observed (H3a) but there was evidence that more positively judged experiences were correlated with stronger perceptions of relational bonds with other participants (H3b). This correlation was also found to persist when a second measure of relational bonds was collected following a three-day delay. Study 2 also failed to detect any significant correlation with fusion in either session and furthermore the results from

additional regressions contradicted the hypothesised connection between subjective positive assessment of ritual experiences and relational bonds.

In summary, contrary to the findings from the martial arts study the experimental studies reported mixed results in Study 1 and null results in Study 2. Finally, the firewalking data did not speak directly to the proposed connection due to a lack of relevant measures addressing subjective affective responses to the ritual experience.

Table 5-4: Summary of Findings for Core Q4

Q4: Are the individual and social effects generated from collective dysphoric rituals distinct from those of collective, high-arousal euphoric rituals?				
	BJJ Promotions	Artificial Ritual Study 1	Artificial Ritual Study 2	Shinto Firewalking
Overall	Positive	Mixed	Mixed	--
Summary of Findings	- Positive subjective experiences associated with identity fusion and relational bonds.	- Positive subjective affect associated with relational bonds but not identity fusion.	- Stronger implicit preference for categorical group members invoked in shared dysphoria condition. However, greater preference for relational group members demonstrated in euphoric condition. - No association with positive subjective affect and fusion or relational bonds.	--

Methodologically the thesis had two aims: first, to develop and assess the effectiveness of the various novel research methodologies applied and second, to demonstrate the utility of employing multiple methodologies across different contexts in order to gain broader insight and provide independent tests of the reliability of findings.

On the first aim, the novel artificial ritual protocols reported in chapter 3 represent a clear improvement over previous efforts (Richert et al., 2005) in that they have better verified affective manipulations, employed a much larger number of independent groups,

and incorporated a wide variety of methods to reduce potential for researcher bias to influence results. The second version of the artificial ritual also included experimenter blinding procedures that enabled researchers collecting key outcome measures to remain blind to condition. Furthermore, significant efforts were made to utilise an actual collective ritual performance as opposed to a more abstract ritual-like task (Legare & Wen, 2014). The utilisation of different behavioural outcomes between the two studies reported made comparisons between studies less precise but also served to illustrate the flexibility of the core methodology. It is anticipated that through refinements to the core framework established further improvements will be possible that will enable the protocol to be applied in a variety of experimental contexts.

There were, however, a number of issues with both of the current designs, the primary one being the poor performance of the physiological HR measures. In large part this seemed to be a result of displacement of chest sensors occurring due to the amount of movement required in the ritual task. However, HR monitors are used extensively in research that requires participants to perform intense physical activity and therefore it may be that the particular equipment employed was not of sufficient quality to provide accurate measurements. Any future studies should conduct extensive pre-testing of physiological measures to attempt to overcome this limitation.

The poor performance of the HR recording equipment is not the only point of concern with this aspect of the study. Arguably a greater issue is the salience that participants attached to the measures in the SER responses. It was clear from qualitative evaluation of responses in both studies that fixing the HR equipment was considered a central

component of many respondent's experimental session, despite efforts being made to reduce the time and emphasis placed on this aspect of the studies. Using more accurate recording devices will not resolve this issue and instead is likely to exacerbate it if the fitting proves more invasive. One solution may be to employ a fictional dual-study cover story, akin to that used in Study 2, wherein HR monitors are used for some alternative purpose in a disguising Study 1 and then left fitted for Study 2 the ritual study.

Of further concern is the lack of ability across both studies (in piloting Study 1 and through manipulation checks in Study 2) to distinguish between euphoric and neutral affective manipulations. Pre-testing of manipulations failed to resolve this issue in Study 2 and thus, it was not possible to identify clear significant differences between the euphoric and neutral conditions. There are various well validated positive affective manipulations employed in psychology, but the majority of these are ill fitted to the ritual protocol, as it currently exists. One promising avenue that has not been explored fully in the current design is the use of validated affective imagery. Using validated affect inducing images, potentially from a source such as the IAPS (P. J. Lang et al., 2008), that are introduced as representative aspects of some original ritual context could serve to further enhance the strength of the manipulation. Alternatively, a different category of audio stimulus that offers clearer distinctions between positive and neutral variations, such as adult group laughter (euphoric), inaudible dialogue (neutral), and adult screams (dysphoric) might help to resolve these limitations.

In regards to the quasi-experimental field study, the use of pre- and post-event time periods is not revolutionary but it remains rarely applied in studies of rituals and the

results reported in chapter 4 suggest that this is unfortunate as there were significant changes detected between periods. For instance, in the response to targets of fusion between observers and performers. The reliability of findings could also be greatly improved if it was possible to arrange for within-group pre- and post-measures to be collected. This would require significant coordination with relevant festival or ritual authorities and would also introduce bias in responses due to the fatigue of repeated measures but the potential gains in reliability and validity seem to make this a worthwhile avenue to pursue.

For the second methodological goal, the thesis seems to have served its purpose admirably. A number of results are clearly distinct between studies and this not only cautions against unwarranted generalisations but also provides fertile ground for further research efforts. For instance, although the experimental lab studies proved poorly equipped to generate any significant levels of identity fusion amongst participants, an experimental priming methodology using recordings of ritual performances with relevant ingroup members might be capable of partly reproducing the shifts in fusion observed in the firewalking fieldwork data. Whether such shifts differed for those with experience as performers could then be examined.

Avenues for Future Research

At the end of each chapter and in the conclusions above a variety of potential avenues for future research have been identified and discussed. What follows below is thus not intended to offer an overall summary of all of the areas identified in the thesis but rather specific topics and methodologies that seem particularly promising:

Based on the contrasting results reported between Xygalatas et al. (2013b) and this thesis in relation to prosociality and inclusive identification following extreme ritual events, there is a need to examine these issues at further field sites. It may be the case that Shinto festivals and the unique religious environment in Japan are an outlier but at present there is no way to determine this. Future studies should thus address the roles played by religious devotion and cultural norms in determining responses to extreme rituals.

There is significant scope for priming methodologies to be used to invoke memories of personally consequential dysphoric and euphoric experiences in order to explore proposed correlations with identity fusion and prosociality. Participants could be asked, for instance, to provide accounts of collective dysphoric events and then asked follow up questions to examine their subjective affective response to these experiences. Based on the studies reported in this thesis, particularly the martial arts data, I would predict that many dysphoric events will be classified subjectively as positive experiences that helped an individual to develop personally or deepen their bonds with others who shared in the event. Whether these subjective responses prove more predictive of prosocial tendencies, relational bonds, and group affiliation, than the character of the original experiences remains an open question. But even if it proves to be the case that subjective assessment is core to the outcomes in question, it could still be that dysphoric events are particularly effective in generating long lasting and strongly felt positive associations. This could also be examined through a priming methodology by contrasting the subjective assessments and motivational forces generated through recalling collective euphoric and collective dysphoric events.

Next, as the artificial ritual studies demonstrated, generating substantive and personally consequential experiences under laboratory conditions is not an easy endeavour. Indeed, when dealing with dysphoria there are also clear ethical issues with attempting to produce experiences that mimic the terror or physical stress that often accompany extreme rituals. A plausible alternative, which is currently being pursued productively by a number of researchers, is to bring the lab ‘into the field’ and make use of the dramatic, painful, and personally consequential rituals that people participate in voluntarily across the world every day. These efforts should be encouraged and this thesis has offered its own contribution through quasi-experimental methodology employed in the Shinto firewalking research. But there is another alternative that to date has remained underutilised in CSR research. That is, bringing the field into the lab, not with naïve students performing unfamiliar rituals or ritual-like tasks, but rather through conducting studies with experienced ritual performers and their associated communities. For example, a potential follow-up study to the results reported in Chapter 4 could be to invite firewalking performers from Shakotan and Furubira to attend a mobile ‘lab’ at which a between-subjects priming methodology, similar to the one discussed above, could be used to invoke salient, personally consequential experiences and then subjects tested on perceptions of group identification, fusion, and prosociality.

Concluding remarks

The primary goal of this thesis was to contribute to ongoing efforts within the CSR to empirically investigate rituals and their personal and social effects. The studies reported have achieved this providing additional evidence that strengthen claims concerning a link

between rituals and their ability to induce and promote social bonding but call into question other proposals, such as a generalisable association between extreme rituals and prosociality and the strength of the connection between dysphoric experiences and identity fusion. Taken collectively the findings reported thus make a substantial contribution to our knowledge about collective dysphoric rituals and their affiliative consequences.

Yet, the results reported ultimately only address one of the hypothesised outcomes of dysphoric rituals. This thesis did not, for instance, directly address the substantial issues that still surround dysphoric rituals and their role in episodic mnemonic formation, nor did it distinguish clearly between the dysphoric impacts derived from physical pain and psychological distress. The point here is not to offer a case against the benefits of making use of the fractionating strategy currently endorsed by the majority of CSR researchers but to emphasise that such efforts ideally should be accompanied by efforts to incorporate findings into larger theories that synthesize a broad range of findings and stimulate further research. A fractionating strategy does not automatically entail the denigration of broader synthesising theories but in practice this can occur, since wider theories tend to be stronger magnets for criticism and attract counter-examples. In short, while I strongly endorse the view that it is crucial to make appropriately restricted generalisations, I also believe that this should not discourage researchers from seeking to place their findings within broader theoretical frameworks and discussions. Relatedly, and as discussed in Chapter 1, I reject the arguments that ritual and religion are too artificial and culturally specific categories to serve as meaningful and coherent analytical categories and instead agree with those who

find them useful and suggest that they relate to some identifiable and cross-culturally recurrent phenomena.

Consequently, as the previous chapters demonstrate, the current thesis has derived its hypotheses and sought to situate its findings in relation to broader theories of ritual and religion, particularly in regards to Modes theory and the ever shifting constellation of costly signalling theories. The studies and results reported in each chapter are intended to contribute to the refinement of these theoretical models by extending their relevance and simultaneously recognising their limitations. Although it is essential for any genuine empirical effort that key theories are capable of falsification, a willingness to substantially revise theoretical models in the face of new evidence should not be regarded as a problematic practice. Indeed, a willingness to update theories on the basis of new evidence is actually a hallmark of good science.

There is clearly much work left to be done and although progress can be frustratingly slow and painful at times, there is the promise of a collective good that can be achieved through developing a better understanding of the rich world of rituals that surround us all.

APPENDICES

APPENDIX I. BJJ Questionnaire

Note: the contents of this questionnaire were presented in a different format online with branching logic and certain pop up texts. The current summary is provided only as an additional reference.

Thanks for taking part! First, some general questions about your training history. If you cannot answer any of the questions just make a rough estimate or indicate that the question is not applicable to you by ticking or typing 'N/A'. If you train with more than one BJJ school or travel between different schools, then for questions asking about your specific BJJ school please answer with regards the group you currently feel the strongest association with. Similarly, if you are currently not training then please answer in regards to the group you previously felt the strongest association with.

Also please note throughout the questionnaire that the term 'BJJ school' is intended to act as a collective term to refer to all BJJ academies/gyms/schools/teams etc.

1. Approximately how long in total have you been training in BJJ? (Please answer using years & months)
2. Have you had any sustained breaks in your training?
3. Approximately how long did the training breaks last collectively?
(Please answer using years and months)
4. Are you currently training?
5. Approximately how long have you been training with your current BJJ school? (Please answer using years & months)
6. What is your current belt in BJJ?
7. Does your school use stripes as well as coloured belts to differentiate training achievements?
8. How many stripes do you have?
9. Which of the following most accurately describes your role in your BJJ school:
(Select one or more boxes)
Head Instructor (1)
Assistant Instructor (2)
Owner/Organiser (3)
Professional Competitor (4)
Amateur Competitor (5)
Non-Competing Student (6)
Complete Beginner (7)
10. On an average week, approximately how many hours do you train in BJJ?
(Please answer in hours)

11. Do you, or have you in the past, competed in BJJ tournaments?
12. On average, how many competitions per year do you currently take part in, as a competitor:
13. On average, how many competitions per year do you currently attend, as an observer:
14. What are your motivations for training in BJJ? (Use your cursor to drag the answers supplied into rank order with 1 signifying the most important motivation and 6 the least important)
 - _____ To compete in tournaments (1)
 - _____ To learn self-defence (2)
 - _____ To train for fitness (3)
 - _____ To become an instructor (4)
 - _____ To compete in Mixed Martial arts (5)
 - _____ To socialise (6)
 - _____ For fun/enjoyment (7)
15. Are there any other important factors not included above which motivate your training in BJJ?
Please briefly describe any other significant factors which motivate you to train.
16. Do you train with any other martial arts groups?
17. How many other groups do you train with?
18. What, if any, affiliation does your BJJ school have?

Section 2 of 4: Half way there... This section is focused on your experience with your experiences at belt promotions/gradings. If you cannot answer any of the questions exactly, either select the most applicable answer or tick/type 'N/A'.

Have you attended any belt promotions/gradings during your time training in BJJ?

The questions below concern specific details of your BJJ school's belt promotion/grading ceremonies. The questions mainly ask about the formal belt promotion/grading practices at your current, or most recent, BJJ school, however please read the questions carefully as some refer to your wider experiences and informal situations.

1. Approximately how many formal belt promotions/gradings have you attended throughout your entire training history (including as an observer)?
2. Approximately how many informal belt promotions/gradings (i.e. those which occur spontaneously as part of/at the end of a normal class) have you witnessed throughout your entire training history (including as an observer)?
3. On average, how often does your school hold formal belt promotions/gradings?
4. Have all the BJJ belt promotions/gradings you attended been with the school you are currently training with?
5. Approximately how many belt promotions/gradings have you attended at your current school?

6. Do the belt promotions/gradings at your current, or most recent, BJJ school involve a belt whipping gauntlet?
7. At this school approximately, how many times have you received a belt whipping as part of your belt promotion/grading?
8. At this school approximately, how many times have you inflicted a belt whipping as part of other members' belt promotions/gradings?
9. Do belt whipping gauntlets at your current BJJ school require that the person receiving the whipping remove their gi jacket?
10. Approximately how long does it take one individual to complete a belt whipping gauntlet at your school?
11. On average, who is involved in inflicting whips during a single belt whipping gauntlet at your BJJ school? (Tick as many boxes as appropriate)
 The instructor(s) (1)
 Higher Grades (Brown & Purple Belts) (2)
 Lower Grades (Blue Belts & Experienced White Belts) (3)
 Beginners (4)
 All of the above / Everyone present (5)
 Other (9) _____
12. How many people on average take part in a belt whipping gauntlet at your BJJ school?
13. Do belt promotions/gradings at your current, or most recent, BJJ school involve being thrown?
14. On average, who performs the throws? (Tick as many boxes as appropriate)
 The instructor(s) (1)
 Higher Grades (Brown & Purple Belts) (2)
 Lower Grades (Blue Belts & Experienced White Belts) (3)
 Beginners (4)
 All of the above / Everyone present (5)
 Other (9) _____
15. Approximately, how many times have you been at a grading ceremony and received throws as part of your belt promotion/grading?
16. Approximately, how many times have you performed throws on someone else as part of their belt promotion/grading ceremony?
17. Do belt promotions/gradings at your current, or most recent, BJJ school involve continuous sparring against multiple opponents (as part of the grading rather than as a normal lesson)?
18. At belt promotions/gradings during which you have been promoted how many have involved sparring continuously against multiple opponents?
19. Approximately how many times have you served as one of a series of sparring opponents during someone else's belt promotions/gradings?
20. On average, who takes part in sparring against those being promoted/graded? (Tick as many boxes as appropriate)
 The instructor(s) (1)
 Higher Grades (Brown & Purple Belts) (2)

Lower Grades (Blue Belts & Experienced White Belts) (3)

Beginners (4)

All of the above / Everyone present (5)

Other (9) _____

21. On average how long in total would you estimate that a belt promotion/grading takes for one person at your BJJ school? (Please answer in mins)

22. Are formal tests of technique a part of your BJJ school's belt promotions/gradings?

23. How long has it been since you last attended a belt promotion/grading ceremony? (Please answer in years/months)

24. How long has it been since you last attended a belt promotion/grading ceremony during which you were promoted? (Please answer in years/months)

25. Approximately how long ago was the first belt promotion/grading ceremony you attended? (Please answer in years/months)

26. Approximately how long ago was the first belt promotion/grading ceremony you attended during which you were promoted? (Please answer in years/months)

27. Before joining your BJJ school how aware were you of your school's belt promotion/grading practices?

I had no idea (1)

I had a vague idea (2)

I didn't know about my school specifically but I knew about BJJ gradings in general (3)

I knew quite a bit about gradings at the school (4)

I had attended/observed gradings at the school before joining (5)

28. Does your BJJ school charge for belt promotions/gradings?

29. How much on average does a belt promotion/grading cost at your BJJ school?

30. Are there any other important elements of a grading at your BJJ school that you think are not mentioned in the above options? (If so, please specify)

Section 3 of 4: This section addresses your personal response to your belt promotion/grading experiences with your current, or most recent, BJJ school. If you have not yet been promoted at a grading, then please answer based on your observations of the belt promotion/grading ceremonies you have witnessed at your BJJ school. If you cannot answer any of the questions either select the most applicable answer or tick/type 'N/A'.

1. How intense would you consider your belt promotion/grading experiences with your current, or most recent, BJJ school?

Not intense at all (1) - Extremely Intense (6)

2. How painful would you consider your belt promotion/grading experiences with your current, or most recent, BJJ school?

Not painful at all (1) - Extremely painful (6)

3. How unpleasant would you consider your belt promotion/grading experiences with your current, or most recent BJJ school?

Not unpleasant at all (1) - Extremely painful (6)

4. How enjoyable would you consider your belt promotion/grading experiences with your current, or most recent, BJJ school?

Not enjoyable at all (1) – Extremely enjoyable (6)

5. How meaningful would you consider your belt promotion/grading experiences with your current, or most recent, BJJ school?

Not meaningful at all (1) - Extremely meaningful (6)

6. How understandable would you consider your belt promotion/grading experiences with your current, or most recent, BJJ school?

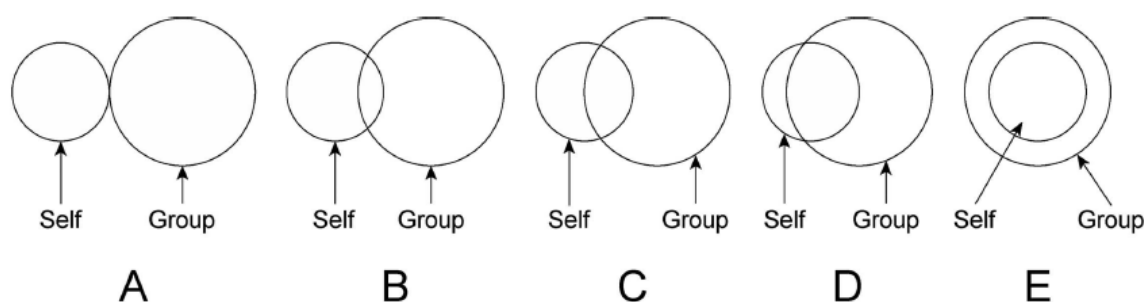
Not understandable at all (1) - Extremely understandable (6)

7. How valuable would you consider your belt promotion/grading experiences with your current, or most recent, BJJ school?

Not valuable at all (1) - Extremely valuable (6)

We are interested in your intuitions about the meaning of the BJJ belt promotion/grading ceremonies you have participated in or observed. In the space provided below, please write down your thoughts about the overall meaning of the grading ceremonies and/or the specific elements mentioned in the previous section.

Look at the images below...



Which pictorial representation below most closely reflects your relationship with your BJJ school?

Below is a list of statements about your BJJ school. Please read the statements and indicate on the scale to what extent you would agree with the statements:

*Also please note that 'BJJ school' is intended to act as a collective term to refer to all BJJ clubs/academies/gyms/ teams etc.

Please indicate your agreement with the statements to the left.

Strongly Disagree (1) - Strongly Agree (6)

1. When someone criticizes my BJJ school, it feels like a personal insult.
2. I am very interested in what members of other teams think about my BJJ school.
3. When I talk about my BJJ school, I usually say 'we' rather than 'they'.
4. Successes of my BJJ school are my successes.
5. When someone praises my BJJ school, it feels like a personal compliment.
6. If the media criticized my BJJ school, I would feel embarrassed.

Please indicate your agreement with the statements to the left.

Strongly Disagree (1) - Strongly Agree (6)

1. I know the personality differences among members of my BJJ school.
2. I think all the members of my BJJ school are somehow personally connected to each other.
3. I think all the members of my BJJ school are somehow personally connected to me.
4. If my BJJ school really needed me I would be willing to donate my free time to it.
5. If my BJJ school were threatened, I would be willing to risk my life fighting to defend it.
6. Most people in my BJJ school behave in a similar way.
7. Most people in the group are similar to each other in their values and in preferences.
8. People in other BJJ groups, generally admire my BJJ school.
9. In general, my BJJ school is not respected by other BJJ groups.

Please indicate your agreement with the statements to the left.

1. I am one with my BJJ school.
2. I feel immersed in my BJJ school.
3. I have a deep emotional bond with my BJJ school.
4. My BJJ school is me.
5. I'll do for my BJJ school more than any of the other members would do.
6. I am strong because of my BJJ school.
7. I make my BJJ school strong.

Section 4 of 4: To help us contextualise the answers you provided about the BJJ belt promotion/grading ceremonies this section asks a series of more general questions which are designed to help us identify your personal values.

Now we will ask you some questions about yourself. Please indicate to what extent you agree or disagree with the following statements.

Strongly Disagree (1) - Strongly Agree (6)

1. I would rather depend on myself than others. (HORIDV1)
2. I rely on myself most of the time; I rarely rely on others. (HORIDV2)
3. I often do my own thing. (HORIDV3)
4. I enjoy being unique and different from others in many ways. (HORIDV4)
5. It is important that I do my job better than others. (VERIDV1)

6. Winning is everything. (VERIDV2)
7. Competition is the law of nature. (VERIDV3)
8. When another person does better than I do, I get tense and aroused. (VERIDV4)
9. If a co-worker gets a prize, I would feel proud. (HORCOL1)
10. The well-being of my co-workers is important to me. (HORCOL2)
11. To me, pleasure is spending time with others. (HORCOL3)
12. I feel good when I cooperate with others. (HORCOL4)
13. Parents and children must stay together as much as possible. (VERCOL1)
14. It is my duty to take care of my family, even when I have to sacrifice what I want. (VERCOL2)
15. Family members should stick together, no matter what sacrifices are required. (VERCOL3)
16. I would do what would please my family, even if I detested that activity. (VERCOL4)

Thank you completing the main questionnaire, your participation is greatly appreciated! Before finishing we would like to record a few pieces of demographic information. Please select the appropriate responses and note that any answers you supply here will be kept entirely anonymous.

1. What sex are you?
2. What is your age? (Please answer in years)
3. In which country do you currently live?
4. What is your nationality?
5. What is your ethnicity? (Tick as many as apply)
6. What best describes your religious affiliation? (Tick any that apply)
7. Is English your first language?

After the survey is complete responses can be entered into a random draw to win £20 (approx. \$30), if you choose to participate. This £20 can be kept for personal use, donated to your BJJ school or the amount can be split between you and your BJJ school. If selected as a winner, you will be contacted by email to arrange a suitable method of payment for the amount you designate for personal use (e.g. PayPal) and any amount designated for your BJJ school will be sent to the school anonymously. Please note that the selected winners' identities will also be kept entirely anonymous along with their responses. Please indicate below whether you would like to be included/excluded from the draw.

Please indicate below, if selected, how much of the £20 reward you would like to keep for personal use and how much you would like to donate to your BJJ school. Please ensure the total amount equals £20.

_____ Reward in £ for Personal Use (1)
 _____ Reward in £ to donate to BJJ School (2)

APPENDIX II. Wailing Infant Questionnaire V1

Rituals – Interpretations and Effects

Session 1

Thank you for choosing to participate in Rituals – Interpretations and Effects. In this study, you will perform a ritual with some other participants, following instructions recently compiled by the School of Anthropology & Museum Ethnography, University of Oxford. We are interested in your intuitions about the meaning of this ritual, and also in the physiological, psychological, and behavioural effects of ritual participation more generally.

Before the experiment begins, we would like you to record your pre-ritual heart rate. Please record your starting heart rate in the space provided.

Heart rate

My heart rate *before* participating in the ritual is: _____ beats per minute.

Please proceed to the instructions for the ritual on the next page.

The following instructions have been supplied by the School of Anthropology & Museum Ethnography, University of Oxford, and is the result of some recent field work undertaken by researchers there. As part of their effort to understand the ritual, the anthropologists are interested in how lay people respond to it. The instructions provided below have been translated and simplified for use in this lab. Similarly, the setting in which you will perform the ritual and the items you will be using in it are approximations provided by anthropologists at the University of Oxford.

Ritual Instructions

The ritual consists of a sequence of behaviours, performed by each participant in turn as the others watch in silence. Please remain silent throughout the entire ritual procedure. Please follow the ritual instructions carefully:

1. The participants will sit in a triangle, as directed.
2. The first participant will perform the following sequence of behaviours, as the other two participants watch in silence, keeping their eyes on him/her:
 - a. Pour the powder into the water in the large earthenware bowl.
 - b. Stir this mixture with your right hand, 25 times, going clockwise.
 - c. Grind the seeds in the mortar, using the pestle.
 - d. Pour the ground seeds into the mixture.
 - e. Put the plant material into the mortar.
 - f. Stir the mixture with your right hand, 25 times, going anti-clockwise.
 - g. Immerse your right hand into the mixture, and sprinkle droplets of water from your right hand over the plant material. You may then dry your hands on the towel provided.
 - h. Pour the small bowl of water into the perforated bowl, and wait until the water has stopped filtering through.
 - i. Crush the plant material with the pestle.
 - j. Put the plant material into the mixture.
3. The second participant, sitting to the left of the first participant, will now move to take his/her place. When the necessary items for the ritual have been prepared, the second participant will perform the sequence of behaviours described above, as the others watch in silence, keeping their eyes on him/her.
4. The third participant will now take the place of the second participant. When the necessary items for the ritual have been prepared, the third participant will perform the sequence of behaviours described above, as the other watch in silence, keeping their eyes on him/her.

Please answer the following questions to enable us to look at the effects of ritual participation as well as your intuitions about this particular ritual.

Heart rate

My heart rate *after* participating in the ritual is: _____ beats per minute.

My *average* heart rate during the ritual was: _____ beats per minute.

My *maximum* heart rate during the ritual was: _____ beats per minute.

Mood

This scale consists of a number of words and phrases that describe different feelings and emotions. Read *each* item and then mark the appropriate answer in the space next to that word. Indicate to what extent you are feeling this way *right now* (that is, at this present moment). Use the following scale to record your answers:

1	2	3	4	5
very slightly or not at all	a little	moderately	quite a bit	extremely
ashamed _____	upset _____	interested _____	scared _____	
irritable _____	hostile _____	determined _____	enthusiastic _____	
afraid _____	guilty _____	strong _____	jittery _____	
nervous _____	active _____	excited _____	distressed _____	
alert _____	proud _____	inspired _____	attentive _____	

-

Ritual meaning

We are interested in your intuitions about the meaning of the ritual you just performed. In the space provided below, please write down your thoughts about the meaning of this ritual, what it may have been intended to do and symbolize.

The following questions are about how you feel about your fellow participants. Recognizing the limited amount of time you have spent together do you:

1. Feel like you know any of the other participants well?

Not at all	Just a little	Somewhat	Moderately	Quite a lot	Very much
0	1	2	3	4	5

2. Feel a sense of connection with any of the other participants?

Not at all	Just a little	Somewhat	Moderately	Quite a lot	Very much
0	1	2	3	4	5

3. Feel like you share something meaningful with any of the other participants?

Not at all	Just a little	Somewhat	Moderately	Quite a lot	Very much
0	1	2	3	4	5

4. Feel an emotional attachment to any of the other participants?

Not at all	Just a little	Somewhat	Moderately	Quite a lot	Very much
0	1	2	3	4	5

The following questions are about how one aspect of the ritual made you feel. Please indicate how the *sound of the baby crying* made you feel, using the descriptions and scale provided:

1 very slightly or not at all	2 a little	3 moderately	4 quite a bit	5 extremely
sympathetic _____	soft-hearted _____	distressed _____	uneasy _____	
touched _____	compassionate _____	troubled _____	alarmed _____	

Demographics

Age: _____

Gender (Tick any that apply):

☐ Female

☐ Male

☐ Other

Ethnicity (Tick any that apply):

☐ White (e.g., White British, White Irish)

☐ South Asian (e.g., Indian, Pakistani, Bangladeshi)

☐ Black (e.g., Caribbean, African)

☐ Mixed (e.g., White/Black Caribbean, White/Black African)

☐ Chinese (e.g., British-born Chinese, Mainland Chinese)

☐ Other (Please specify): _____

Nationality (Tick any that apply):

☐ British

☐ English

☐ Scottish

☐ Welsh

☐ Northern Irish

☐ Other (Please specify): _____

Religion (Tick any that apply):

☐ Christian (e.g., Anglican, Roman Catholic, Other Christian denomination)

☐ Buddhist

☐ Hindu

☐ Jewish

☐ Muslim

☐ Sikh

☐ None

☐ Other (Please specify): _____

Is English your first language? (Tick ONE)

☐ Yes

☐ No

Rituals – Interpretations and Effects

Session 2

Instructions

This part of the experiment is about decision making, and will involve a resource management task. You will be paid for participating in this task in addition to the payment you will receive for participating in the experiment in general. The additional amount of money you earn depends on the decisions that you and the other participants make. All the money that you earn after this will be yours to keep, and your earnings will be paid to you in cash at the end of today's session. As with all other parts of this experiment, your name will never be associated with any of your decisions. In order to keep your decisions private, please do not reveal your choices to any other participant.

You will engage in five rounds of decision making. In each round, you and every other member of **Team Ten** will be endowed with 10 “tokens.” You must decide how to divide your tokens between a PRIVATE ACCOUNT and a group TEAM TEN ACCOUNT. Each person in the group has a separate PRIVATE ACCOUNT, but there is only one TEAM TEN ACCOUNT. Each person in the group will be making a decision with the same accounts and payments.

In each round you will earn 20p for each token placed in your PRIVATE ACCOUNT. Thus, if you choose to place all of your tokens in your PRIVATE ACCOUNT you would earn $(10 \times 20p) = £2.00$ in that round.

In each round you and every other member of the group will earn 10p for each token placed in the TEAM TEN ACCOUNT (a total of 30p for the group for each token). All members of the group can place tokens in the TEAM TEN ACCOUNT.

To summarize, you will earn:

20p for each token in your PRIVATE ACCOUNT

+ 10p for each token you place in the TEAM TEN ACCOUNT

+ 10p for each token other members of the group place in the TEAM TEN ACCOUNT

Imagine the following situation and calculate your payoff for that round.

Example

If I place ____5____ tokens in the PRIVATE ACCOUNT and ____5____ tokens in the TEAM TEN ACCOUNT, and the two other people in my group place a total of ____10____ tokens in the TEAM TEN ACCOUNT, I will earn:

_____ for the tokens in my PRIVATE ACCOUNT (20p each)

_____ for the tokens I put in the TEAM TEN ACCOUNT (10p each)

_____ for the tokens the other people in my group put in the TEAM TEN ACCOUNT
(10p each)

TOTAL: £ _____

You will make 5 allocation decisions like the one in the example above. On each round, we will collect your decisions and report the number of tokens in the TEAM TEN ACCOUNT. We will then proceed to the next round.

At the end of the five rounds, we will choose one round on which you will be paid. Thus, you should think carefully about your decision in each round because you do not know which round will determine your payoff. To determine the round for which you will be paid, we will randomly choose a number between one and five by having one of you draw a number from an envelope with numbers from 1 to 5. That number will determine which *one* round you will be paid for. For example, if we draw a 1 you will be paid for your decision on Round 1.

During the experiment, you are not permitted to speak or communicate with the other participants. If you have a question while the experiment is going on, please raise your hand and the experimenter will come to you. If you have any questions at this time, please raise your hand and the experimenter will come to your seat to answer it.

DECISION SHEET

To summarize, you will earn:

$20p \times$ the no. of tokens you keep in your PRIVATE ACCOUNT

$+ 10p \times$ the no. of tokens you invest in the TEAM TEN ACCOUNT

$+ 10p \times$ the no. of tokens the other group members invest in the TEAM TEN account

Round 1.

I place _____ tokens in my PRIVATE ACCOUNT

I place _____ tokens in the TEAM TEN ACCOUNT

(These choices must sum to 10 tokens)

The experimenter will tally the total number of tokens in the
TEAM TEN ACCOUNT: _____

Round 2.

I place _____ tokens in my PRIVATE ACCOUNT

I place _____ tokens in the TEAM TEN ACCOUNT

(These choices must sum to 10 tokens)

The experimenter will tally the total number of tokens in the
TEAM TEN ACCOUNT: _____

Round 3.

I place _____ tokens in my PRIVATE ACCOUNT

I place _____ tokens in the TEAM TEN ACCOUNT

(These choices must sum to 10 tokens)

The experimenter will tally the total number of tokens in the
TEAM TEN ACCOUNT: _____

Round 4.

I place _____ tokens in my PRIVATE ACCOUNT

I place _____ tokens in the TEAM TEN ACCOUNT

(These choices must sum to 10 tokens)

The experimenter will tally the total number of tokens in the
TEAM TEN ACCOUNT: _____

APPENDIX III. Wailing Infant V2 Video Instructions / Ritual Sequence

Note: the contents of all this material presented were translated, edited, and appeared in a different layout in the experimental material. The summary below is provided only for additional reference.

SCREEN 1

This video will display the sequence for a ritual task, which you will shortly be asked to copy and perform. Please pay attention and try to remember the sequence of events.

Participants should:

1. Not talk or communicate directly during the ritual performance.
2. Treat the ritual performance seriously and perform the steps respectfully.
3. Try to remember as much of the sequence as possible and only refer to the written instructions when absolutely necessary.
4. Remember that the main ritual sequence will be repeated three times, as each participant will perform each of the different roles once and that there is an additional final step after the third repetition.
5. The colour of the powder changes during each repetition. Yellow for the 1st sequence, Red for the 2nd sequence and Blue for the 3rd sequence.

SCREEN 2

This video was filmed to provide instruction, when you perform the task certain additional objects or the environmental conditions (such as lighting & sound) may be altered to more accurately simulate the original ritual context. The sequence performed should remain the same and you should only use the items shown in this video.

Please note the ritual requires three people to perform the necessary roles:

1. Timekeeper (Azar) – Person to the Left
2. Purifier (Reza)- Person to the Right
3. Performer (Zaritto)- Person at the Center in front of the Altar

These roles are rotated in a clockwise direction and the main sequence is repeated 3 times before the final stage of the ritual.

RITUAL SEQUENCE

Everyone: Sit on the white circles. Pick up and put on the necklaces.

Timekeeper: Ring the bell once

Purifier:

1. Place one spoonful of coloured powder into the smaller wooden bowl on the altar.
2. Use the empty jar & pour water into the large wooden bowl in the centre of the altar.
3. Put one spoonful of seeds into the mortar (grinder) base.
4. Select one of the plant leaves and place it on the curved holder on the altar.

Stage 2: Central Performance

Timekeeper: Ring the bell twice

Performer:

1. Pour the yellow (red for 2nd repetition, blue for 3rd) powder into the water in the larger wooden bowl.
2. Use the stick to stir the water 25 times in a clockwise direction and then put it back on its holder. While doing so stare intently at the face of the statue.
3. Grind the seeds carefully in the mortar, using your hand to prevent the seeds from spilling, then transfer them to the mixture in the large wooden bowl.
4. Pick up the plant material, raise slowly to your face then blow gently on it and place it in the mortar.
5. Grind the plant material thoroughly and then transfer the ground plant into the mixture in the large wooden bowl.
6. Use the stick to stir the water 25 times in an anti-clockwise direction and then put it back on its holder. While doing so stare intently at the face of the statue.
7. Transfer the mixture to the collective bowl in front of the altar marked with string.

Stage 3: Purification

Timekeeper: Ring the bell three times

Purifier:

1. Clean the small wooden bowl with the towel provided
2. Clean & Dry the wooden bowl used for mixing with the towels provided
3. Clean the mortar (grinder) with the towel provided

Stage 4: Change Roles

Timekeeper: Clap the wooden blocks together loudly (once)

Everyone: Remove the necklaces and place on the wooden holder & rotate positions in a clockwise direction

- (REPEAT STAGES 1-4) in new positions
- (REPEAT STAGES 1-3) in new positions but finish by completing Stage 5

Stage 5 Conclude Ritual (after Timekeeper claps on third sequence)

Everyone: Reach over, place a hand on the statue's body and gently place it into the centre of the mixture in the collective bowl which contains the coloured mixture.

Everyone: Starting with the performer, use the stick (held in your right hand) to sprinkle a few drops from the mixture in the collective bowl onto the head of the statue. Pass the stick in a clockwise direction using both hands.

Timekeeper: Clap the wooden blocks together loudly (twice)

You will shortly perform the ritual so finally please remember:

1. Mistakes are quite common during replications and you should not be concerned if you realise you have made an error or forgotten a step. If this happens simply move on and try to complete the remaining steps.
2. The ritual performance should be treated seriously and respectfully. No talking is permitted during the ritual performance and the experimenter will not intervene once the ritual begins, unless absolutely necessary. PPTs should exchange positions without waiting for instructions from the experimenter and without verbally communicating with the other PPTs.
3. To make the ritual setting more accurately simulate the original context, various environmental factors (such as sound & lighting) will be altered based on the instructions provided by the archaeologists. The ritual should be understood as beginning when the experimenter says 'please begin' and starts playing the background sound.
4. When not performing an action, participants must focus intently on whatever action is being performed. Please do not lose attention or focus.

RITUAL SEQUENCE – ENGLISH VERSION

Stage 1: Preparation

Everyone: Sit on the white circles. Pick up and put on the necklaces.

Timekeeper: Ring the bell once  x 1

- Purifier:**
1. Place one spoonful of powder into the smaller wooden bowl on the altar. 
 2. Use the empty jar & pour water into the large wooden bowl in the centre of the altar.
 3. Put one spoonful of seeds into the grinder base. 
 4. Select one of the plant leaves and place it on the curved holder on the altar. 

Stage 2: Central Performance

Timekeeper: Ring the bell twice  x 2

- Performer:**
1. Pour the yellow powder into the water in the larger wooden bowl. 
 2. Use the stick to stir the water 25 times in a clockwise direction and then put it back on its holder. While doing so stare intently at the face of the statue.  x 25 
 3. Grind the seeds carefully in the mortar, using your hand to prevent the seeds from spilling, then transfer them to the mixture in the large wooden bowl.  
 4. Pick up the plant material, raise slowly to your face then blow gently on it and place it in the grinder. 
 5. Grind the plant material thoroughly and then transfer the ground plant into the mixture in the large wooden bowl.  
 6. Use the stick to stir the water 25 times in an anti-clockwise direction and then put it back on its holder. While doing so stare intently at the face of the statue.  x 25 
 7. Transfer the mixture to the collective bowl in front of the altar marked with string. 

Stage 3: Purification

Timekeeper: Ring the bell three times  x 3

- Purifier:**
1. Clean the small wooden bowl with the towel provided
 2. Clean & Dry the small wooden bowl with the towels provided
 3. Clean the grinder with the towel provided



Stage 4: Change Roles

Timekeeper: Clap the wooden blocks together loudly (once)  x 1

Everyone: Remove the necklaces and place on the wooden holder & rotate positions in a clockwise direction

- (REPEAT STAGES 1-4) in new positions

- (REPEAT STAGES 1-3) in new positions but now

Stage 4: Conclude Ritual (after Timekeeper claps on third sequence)

Everyone: Reach over, place a hand on the statue's body and gently place it into the centre of the mixture in the collective bowl.



Everyone: Starting with the performer, use the stick (held in your right hand) to sprinkle a few drops from the mixture in the collective bowl onto the head of the statue. Pass the stick in a clockwise direction using both hands.

Timekeeper: Clap the wooden blocks together loudly (twice)  x 2

儀式遂行の手順

ステージ1 | 準備

■全員

丸杵の席に座り、ネックレスを首にかける。

■タイムキーパー

ベルを1回鳴らす。

■お清め担当

1. ひとさじ分の色のついた粉末を、祭壇中央にある小さい木製の器に入れる。
2. 空のビンを使って、祭壇中央にある大きい木製の器に水を注ぐ。
3. 種をひとさじ分すくい取り、石製のすり鉢に入れる。
4. 植物の葉を1枚選び取り、祭壇中央の竹の皿の上に置く。



6. 木の枝を使って反時計回りに25回水をかき混ぜ、木の棒を皿の上に戻す。
この作業をしている間は、木製の像の顔をじっと見つめる。
7. 大きい方の木製の器に入った水を、儀式場中央にあるひもの付いた器の中に移す。

ステージ3 | お清め

■タイムキーパー

ベルを3回鳴らす。

■お清め担当

1. 乾いたタオルで小さい木製の器をいきれいにする。
2. 濡れたタオルと乾いたタオルで大きい木製の器をいきれいにする。
3. 乾いたタオルで石製のすり鉢を拭いてきれいにする。

ステージ4 | 役割の変更

■タイムキーパー

木の板を1回叩いて、大きな音を出す。

■全員

1. ネckレスを首から外して木製の皿に置き、席を時計回りにひとつずつずれて、座る位置を交替する。
2. 役割を交代して、ステージ1～4を繰り返す(2回目)。
3. 3回目の儀式はステージ1から3までを行う。

ステージ2 | 儀式(3回繰り返す)

■タイムキーパー

ベルを2回鳴らす。

■儀式遂行者

1. 色のついた粉末を水の入った大きい木製の器の中に入れる。
2. 木の棒を使って時計回り水を25回かき混ぜ、木の棒を皿に戻す。
この作業をしている間は、木製の像の顔をじっと見つめる。
3. 種をこぼさないように手を使いながら慎重に石製のすり鉢ですり潰す。
その後、潰した種を大きい木製の器の中に移す。
4. 葉を取り上げてゆっくりと自分の顔に近づけて息をそっと吹きかけるづけど。
その後、その葉を石製のすり鉢に入れる。
5. 葉っぱをすり潰し、それを大きい木製の器に入れる。



ステージ5 | 終了の儀式

■タイムキーパー

木の板を1回叩いて、大きな音を出す。

■全員

1. 木製の像の胴体部分を掴み、儀式場中央にある、ヒモの付いた器の中にそっと置くづれて。
2. 儀式遂行者は、木の棒を右手で持って器の中に入った水を像の頭の上に数滴振りかける。
3. 自分の左に座っている人に両手で木の棒を渡す。

■タイムキーパー

木の板を2回叩いて、大きな音を出す。

APPENDIX V. Wailing Infant V2 B-IAT Instructions

We are also interested in how completing the ritual task affects your ability to correctly identify names associated with the tribe whose ritual you just performed. As such, you will be shortly provided with instructions and asked to perform a categorizing task of tribal names/titles.

There is an easy way to identify members of the Zanshii tribe as all of their names contain the sound ZA somewhere and similarly all members of the Kokiar tribe contain the sound KO.

On the next page are a list of 8 names from Zanshii tribe and 8 names from the Kokiar tribe which will be used in the categorizing task. You will now have one minute to learn the names before the categorizing task begins.

Please try and concentrate on the names provided and remember the simple categorizing rule:

- Names which include ZA are from the Zanthie tribe
- Name which include KO are from the Kokiar tribe

/ 1 = "Put your middle or index fingers on the E and I keys of your keyboard. Words representing the categories at the top will appear one-by-one in the middle of the screen. When the item belongs to a category on the left, press the E key; when the item belongs to a category on the right, press the I key. Items belong to only one category. If you make an error, an X will appear - fix the error by hitting the other key.

This is a timed sorting task. GO AS FAST AS YOU CAN while making as few mistakes as possible. Going too slow or making too many errors will result in an uninterpretable score. This task will take about 5 minutes to complete."

/ 2 = "See above, the categories have changed. the items for sorting have changed as well. The rules, however, are the same.

When the item belongs to a category on the left, press the E key; when the item belongs to a category on the right, press the I key. Items belong to only one category. An X appears after an error - fix the error by hitting the other key. GO AS FAST AS YOU CAN."

/ 3 = "See above, the four categories you saw separately now appear together. Remember, each item belongs to only one group. For example, if the categories flower and good appeared on the separate sides above - words meaning flower would go in the flower category, not the good category.

The green and white labels and items may help to identify the appropriate category. Use the E and I keys to categorize items into four groups left and right, and correct errors by hitting the other key."

/ 4 = "Sort the same four categories again. Remember to go as fast as you can while making as few mistakes as possible.

The green and white labels and items may help to identify the appropriate category. Use the E and I keys to categorize items into the four groups left and right, and correct errors by hitting the other key."

/ 5 = "Notice above, there are only two cateogries and they have switched positions. The concept that was previously on the left is now on the right, and the concept that was on the right is now on the left. Practice this new configuration.

Use the E and I keys to catgorize items left and right, and correct errors by hitting the other key."

/ 6 = "See above, the four categories now appear together in a new configuration. Remember, each item belongs to only one group.

The green and white labels and items may help to identify the appropriate category. Use the E and I keys to categorize items into the four groups left and right, and correct errors by hitting the other key."

/ 7 = "Sort the same four categories again. Remember to go as fast as you can while making as few mistakes as possible.

The green and white labels and items may help to identify the appropriate category. Use the E and I keys to categorize items into the four groups left and right, and correct errors by hitting the other key."

APPENDIX VI. Wailing Infant V2 Tajfel Allocation Tasks

Welcome to the first session of this 2 week study on "Group Experiences and Recall"

This study will be conducted over two weeks and consists of two parts:

In the first part of the study: you will be asked some questions about your feelings towards different groups to which you belong. Some of the questions will refer to long term groups and some will refer to short term groups including the groups you were allocated to during the previous experiment.

In the second part of the study: most participants will complete a series of memory tasks which involve memorizing a list of items or details from short stories and then answering questions related to these. You can earn bonus payments based on your performance in these tasks.

However, one participant from each group (selected at random) will not perform the memory tasks. Instead, they will complete a series of resource allocation tasks. If you are selected for this task you cannot earn any bonus payments yourself and instead will be paid a fixed bonus amount which does not change regardless of their choices.

If you have any questions, at any time during the experiment, please press the switch labelled on the wall to call the experimenter.

Please read all instructions carefully and thank you for participating.

ATTENTION:

You have been randomly selected as the participant from the current group to complete the bonus allocation task.

This means that you will not participate in the memory recall tasks and thus will not have the opportunity to earn bonus payments yourself. Instead, as compensation and regardless of your choices in the following allocation task you will be paid a fixed bonus amount. This will be paid in addition to the standard payment for participation.

From the choices you make in the following allocation tasks a random selection will be selected and will have real consequences on the bonus amounts paid to other participants. As such please think about the choices you make but also note that there is no wrong or right answer and any responses you choose will be kept entirely anonymous.

Please do not reveal your role to any of the other participants in the study or discuss what the task involves with other people until after the experiment is completely finished at the end of this semester. This is important as discussing the study with other present or future participants could influence their answers and expectations and thus invalidate the research. We appreciate your cooperation in this matter.

On the next page the instructions for the bonus allocation task are provided, please read them carefully and answer the practice questions before continuing to the actual task. If you have any questions, please contact the experimenter by pressing the button located on the wall.

Bonus Allocation Task

You will now complete an allocation task which involves allocating bonus payments between two group members.

In each of the following rounds you will be presented with a table containing two rows with 13 columns (see below) containing numbers which signify bonus payments (in Yen) to be paid to other participants of the study. Although you will know the group the other participants belong to you will not know the specific identities of the individuals you are allocating bonuses to as every participant has been allocated an anonymous code number.

Example :

表中の数字はボーナスの金額を表しています。

(円)

グループ X のメンバーの74番

グループ Y のメンバーの66番

190	180	170	160	150	140	130	120	110	100	90	80	70
10	30	50	70	90	110	130	150	170	190	210	230	250

The two rows indicate the amounts to be allocated to two different participants and you can select any of the 13 bonus options available. However, you must select the payment in the same column for both participants. For instance, if you choose to allocate 100 Yen to Member 74 of Group X then you would also automatically choose to allocate 190 Yen to Member 66 in Group Y. To select the option you desire you should click once anywhere inside either box within the chosen column.

Example: You could choose this the 100/190 option by clicking anywhere within the shaded box. The final step is to confirm your choice by entering the relevant amounts chosen into the boxes below. Please do this now for the current example:

From the choices you make a random sample will be selected and will have direct consequences to the bonus amounts paid to other participants at the end of the study. There is no way to tell which of the questions will have real consequences so please treat each round seriously and make the choice that you consider best.

You will now complete two more practice rounds before beginning the actual allocation tasks. Please fill in below details of the box you have just chosen.

Matrix Type A:

表中の数字はボーナスの金額を表しています。

(円)

日本人の参加者
の51番
外国人の参加者
の96番

190	180	170	160	150	140	130	120	110	100	90	80	70
10	30	50	70	90	110	130	150	170	190	210	230	250

Matrix Type B:

表中の数字はボーナスの金額を表しています。

(円)

外国人の参加者
の96番
日本人の参加者
の51番

70	80	90	100	110	120	130	140	150	160	170	180	190
10	30	50	70	90	110	130	150	170	190	210	230	250

APPENDIX VII. Shinto Firewalking Questionnaire

Note: the contents of this questionnaire were translated and edited and appeared in a different layout on the page. The summary below is provided only for additional reference.

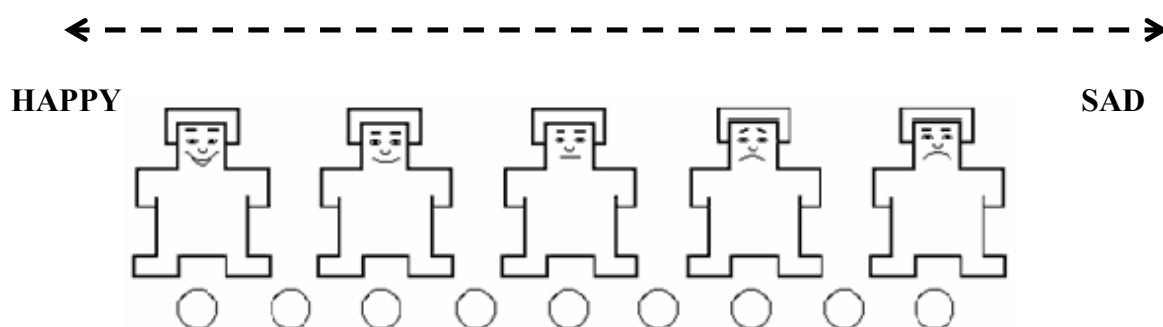
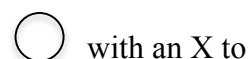
Demographic Information

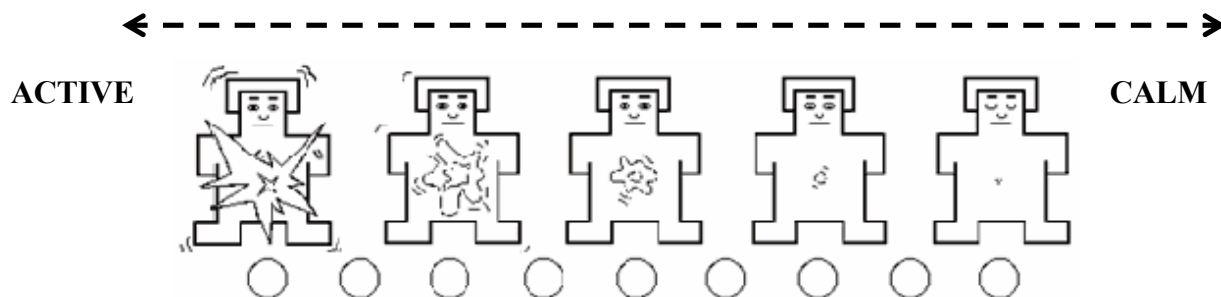
1. What is your sex? Male ☐ Female ☐
 2. How old are you? _____ Years old
 3. How long approximately have you lived in Shakotan? _____ Years
- Or I live somewhere else (please specify): _____
Time living there: _____ Years
4. What is your nationality? Japanese ☐ or Other (please specify):

 5. What is your occupation? _____
 6. What is your relationship status? Single ☐ In a relationship ☐ Married ☐
Other (please specify): _____
 7. How many children do you have? _____

About your Mood

Please use the images and labels as a guide and mark the appropriate indicate how you are feeling *right now*:





About Your Past Festival Experiences

1. How many times have you attended this festival before? _____ Times
2. Have you ever participated in the event as a firewalker before? Yes ☐ No ☐
If yes, how many times a have you been a firewalker? _____ Times

3. What best describes your role in this year's firewalking matsuri?

- ☐ Observer from the local area
- ☐ Observer from outside
- ☐ Volunteer/Organiser
- ☐ Firewalker/Performer
- ☐ Other (please specify) _____

4. From the firewalkers this year do you have any close friends participating?

Yes ☐ No ☐

5. From the firewalkers this year do you have any family members participating?

Yes ☐ No ☐

6. How painful do you think the Firewalking is for the *Tengu*?

Please circle the appropriate number from the choices below.

1	2	3	4	5
Not painful at all	A little painful	Somewhat painful	Moderately painful	Extremely Painful

7. How painful do you think the Firewalking is for the people who carry the *mikoshi*?

Please circle the appropriate number from the choices below.

1	2	3	4	5
Not painful at all	A little painful	Somewhat painful	Moderately painful	Extremely Painful

About The Festival

Please read the following sentences and indicate, *by circling the appropriate number below the sentence*, how far you agree/disagree with the statements’.

1. This matsuri is something that you cannot value in terms of money.

1	2	3	4	5	6	7
Strongly Disagree	Moderately Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Moderately Agree	Strongly Agree

2. This matsuri is something that we should not sacrifice no matter how high the benefits.

1	2	3	4	5	6	7
Strongly Disagree	Moderately Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Moderately Agree	Strongly Agree

3. The firewalking performers are like family to me.

1	2	3	4	5	6	7
Strongly Disagree	Moderately Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Moderately Agree	Strongly Agree

4. This matsuri is very important for *the community of Shakotan*.

1	2	3	4	5	6	7
Strongly Disagree	Moderately Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Moderately Agree	Strongly Agree

5. This matsuri is very important for *me personally*.

1	2	3	4	5	6	7
Strongly Disagree	Moderately Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Moderately Agree	Strongly Agree

6. Participating in this matsuri has helped shape the person I am today.

1	2	3	4	5	6	7
Strongly Disagree	Moderately Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Moderately Agree	Strongly Agree

7. I feel like I share something meaningful with participants in the matsuri.

1	2	3	4	5	6	7
Strongly Disagree	Moderately Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Moderately Agree	Strongly Agree

8. I feel an emotional attachment to participants of the matsuri.

1	2	3	4	5	6	7
Strongly Disagree	Moderately Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Moderately Agree	Strongly Agree

Scenarios

If the shrine were faced with a crisis, which could result in the shrine being closed and the matsuri being cancelled permanently, how many hours would you be willing to donate in an average week to take part in group activities aimed at saving the shrine?

Please answer in hours, using only numbers e.g. 10 = willing to donate 10 hours over one week.

_____ **Hours**

If the shrine requested the community to help advertise/promote the matsuri, how many hours would you be willing to donate in an average week to take part in group activities aimed at promoting the matsuri?

Please answer in hours, using only numbers e.g. 10 = willing to donate 10 hours over one week.

_____ **Hours**

If you could prevent a firewalker from sustaining a severe burn at the matsuri but in so doing would instead put yourself at great risk to sustain a severe burn how likely would you be to act to intervene?

Please circle the appropriate number from the choices below.

1	2	3	4	5
I would definitely not act and avoid any personal risk	I would probably not act and avoid any personal risk	I am not sure if I would or would not act	I would probably act to intervene despite the personal risk	I would definitely act to intervene despite the personal risk

Community

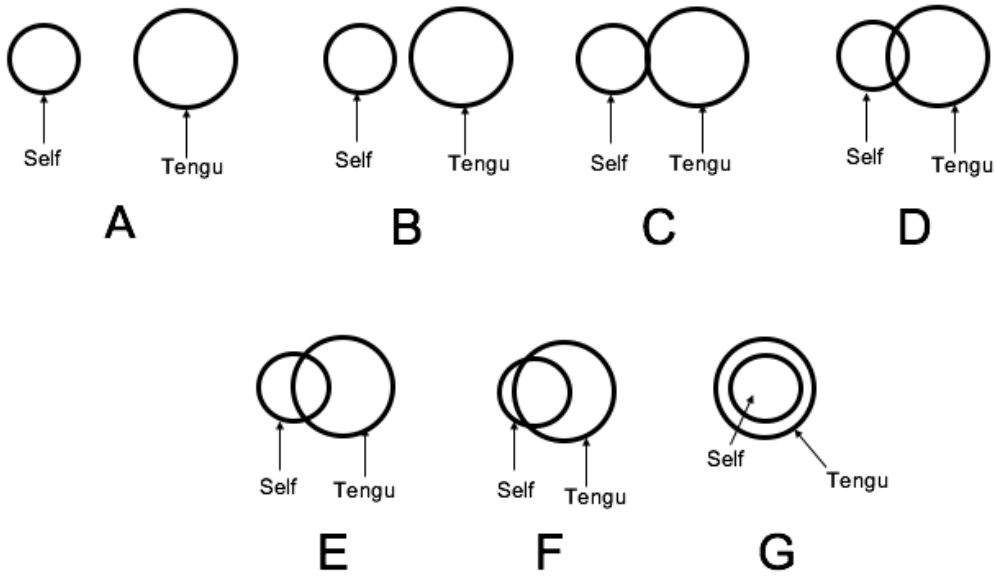
1. Right now, which one of these identities better describes you?

Please mark your answer with an X on the line below.

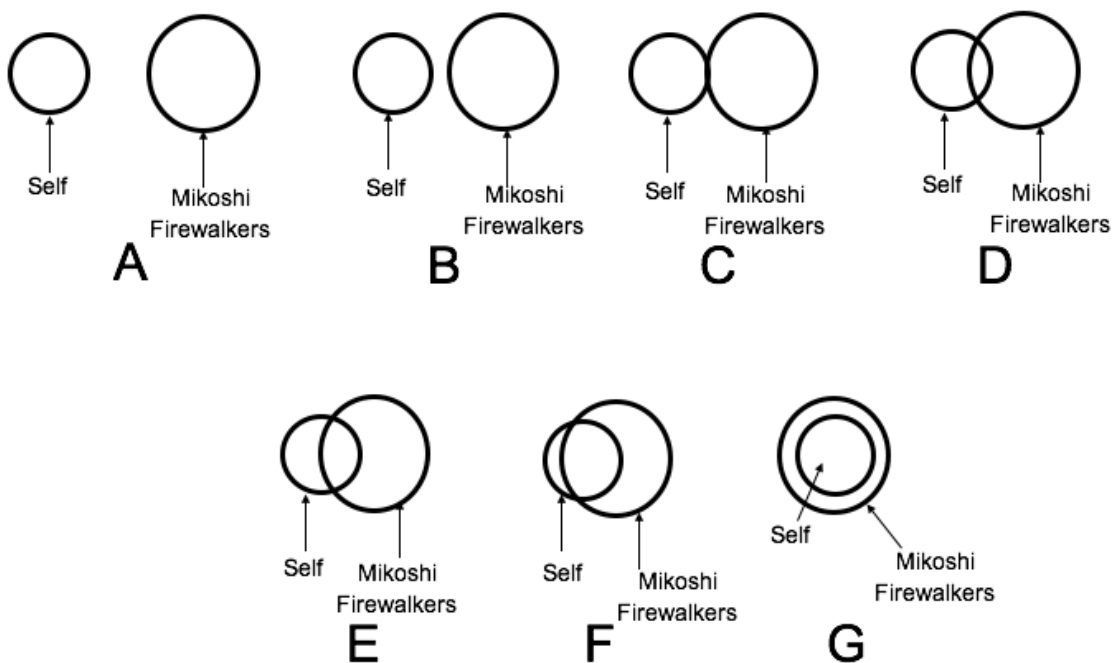
**JAPANESE
PERSON**

**SHAKOTAN
PERSON**

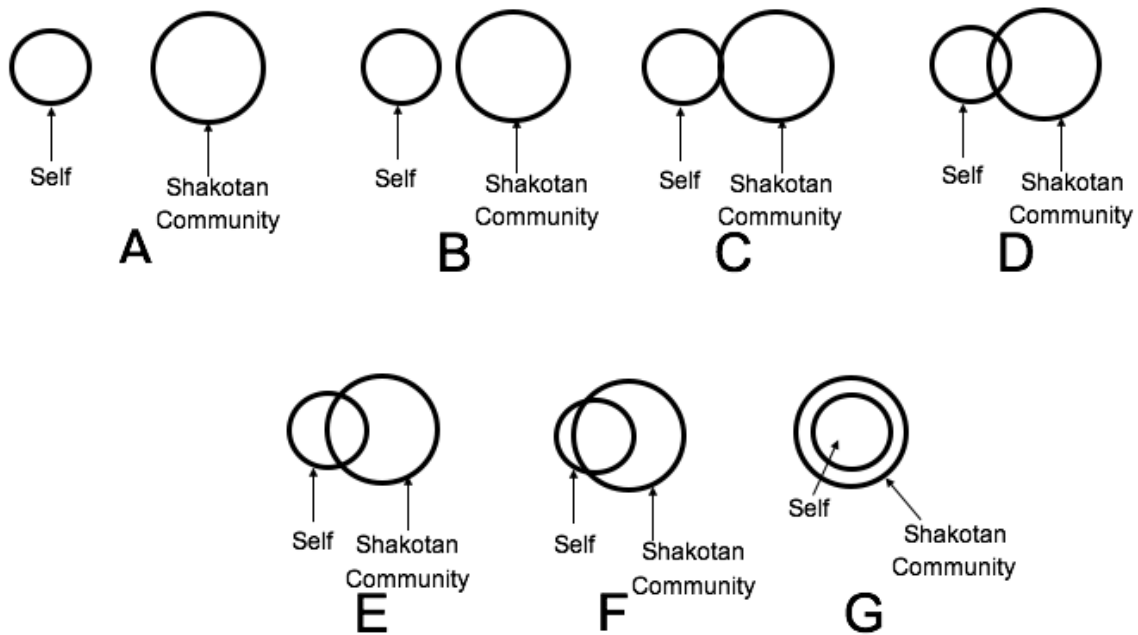
2. Which picture below most closely reflects your relationship with the Tengu?
Please circle the letter below the relevant image.



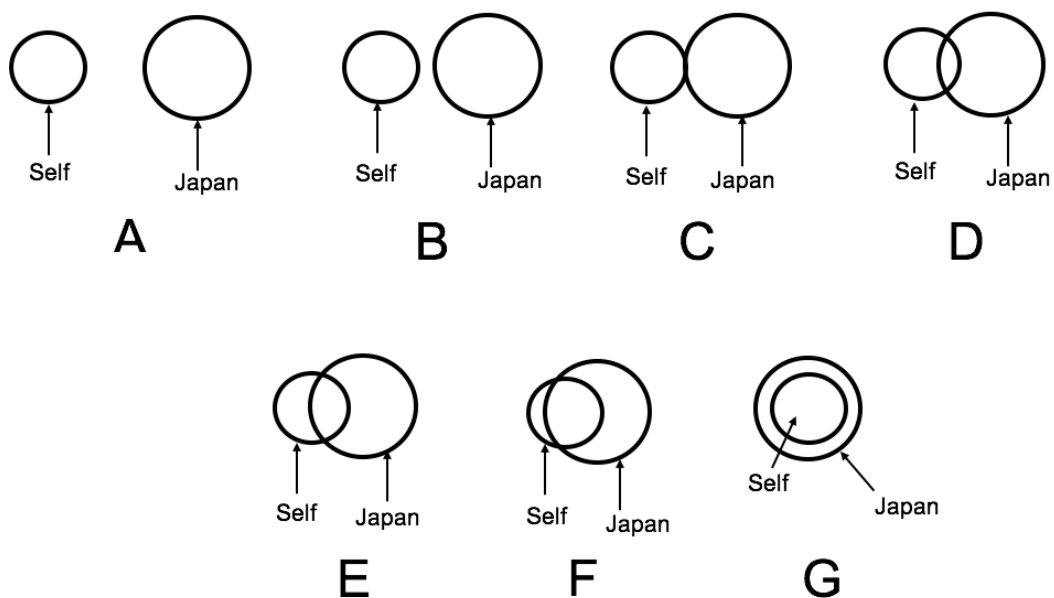
3. Which picture below most closely reflects your relationship with the Mikoshi Firewalkers?
Please circle the letter below the relevant image.



4. Which picture below most closely reflects your relationship with the community of Shakotan?
Please circle the letter below the relevant image.



5. Which picture below most closely reflects your relationship with Japan?
Please circle the letter below the relevant image.



PRIZE LOTTERY

Finally, there is a 15,000 ¥ prize available which will be allocated via a lottery with everyone who completes the questionnaire entered. We expect to hand out 100-200 questionnaires so you have a good chance to win! If you win the prize, you can choose to keep all of the money or if you prefer you can donate some or all of the amount to the shrine. *Your answer will be kept confidential.*

If you win the lottery how would you like to allocate the 15,000 ¥?

Keep Myself: _____ ¥

For the Shrine: _____ ¥

Please note the total must equal 15,000 ¥

BIBLIOGRAPHY

- Abbink, J. G. (1999). Violence, ritual, and reproduction: Culture and context in Surma dueling. *Ethnology*, 28(3), 227–242.
- Abelson, R. P. (1963). Computer simulation of “hot cognition.” In S.S. Tomkins *Computer simulation of personality*. (pp. 277–302), Madison, WI: University of Wisconsin.
- Abrams, D., & Hogg, M. A. (1988). Comments on the motivational status of self-esteem in social identity and intergroup discrimination. *European Journal of Social Psychology*, 18, 317–334.
- Akimoto, S. A., & Sanbonmatsu, D. M. (1999). Differences in Self-Effacing Behavior between European and Japanese Americans Effect on Competence Evaluations. *Journal of Cross-Cultural Psychology*, 30(2), 159–177.
- Alcorta, C. S., & Sosis, R. (2005). Ritual, emotion and sacred symbols: The evolution of religion as an adaptive complex. *Human Nature*, 16(4), 323–359.
- Allen, M. R. (1967). *Male cults and secret initiations in Melanesia*. Melbourne: Melbourne University Press.
- Allison, A. (1994). *Nightwork: Sexuality, Pleasure, and Corporate Masculinity in a Tokyo Hostess Club*. Chicago, IL: University Of Chicago Press.
- Ama, T. (2005). *Why are the Japanese non-religious?: Japanese spirituality: being non-religious in a religious culture*. Lanham, MD: University Press of America.
- Anderson Sawada, J. (1998). Mind and Morality in Nineteenth-Century Japanese Religions: Misogi-Kyo and Maruyama-Kyo. *Philosophy East and West*, 48(1), 108–141.
- Anderson, B. (1983). *Imagined Communities: Reflections on the origin and spread of Nationalism*. London, UK: Verso.
- Anderson, E., McCormack, M., & Lee, H. (2011). Male Team Sport Hazing Initiations in a Culture of Decreasing Homophobia. *Journal of Adolescent Research*, 27(4), 427–448.
- Armstrong, D. F., & Wilcox, S. E. (2007). *The gestural origin of language*. Oxford, UK: Oxford University Press.
- Aronson, E., & Mills, J. (1959). The effect of severity of initiation on liking for a group. *The Journal of Abnormal and Social Psychology*.
- Asad, T. (1993). *Genealogies of Religion: Discipline and Reasons of Power in Christianity and Islam* (Vol. 4). Baltimore, MD: John Hopkins University Press.
- Ashforth, B. E., & Mael, F. (1989). Social identity theory and the organization. *Academy of Management Review*, 14(1), 20–39.
- Atkinson, Q. D., & Whitehouse, H. (2011). The Cultural Morphospace of Ritual Form: Examining modes of religiosity cross-culturally. *Evolution and Human Behavior*, 32(1), 50–62.
- Atran, S., & Ginges, J. (2012). Religious and Sacred Imperatives in Human Conflict. *Science*, 336(6083), 855–857.
- Atran, S., & Henrich, J. (2010). The Evolution of Religion: How Cognitive By-Products, Adaptive Learning Heuristics, Ritual Displays, and Group Competition Generate Deep Commitments to Prosocial Religions. *Biological Theory*, 5(1), 18–30.
- Atran, S., Sheikh, H., & Gomez, A. (2014a). Devoted actors sacrifice for close comrades and sacred cause. *Proceedings of the National Academy of Sciences*, 111(50), 17702–3.
- Atran, S., Sheikh, H., & Gomez, A. (2014b). For Cause and Comrade: Devoted Actors and Willingness to Fight. *Cliodynamics: The Journal of Quantitative History and Cultural Evolution*, 5(1).
- Bachorowski, J.-A., & Owren, M. J. (2001). Not all laughs are alike: Voiced but not unvoiced laughter readily elicits positive affect. *Psychological Science*, 12(3), 252–257.
- Bachorowski, J.-A., Smoski, M. J., & Owren, M. J. (2001). The acoustic features of human laughter. *The Journal of the Acoustical Society of America*, 110(3), 1581–1597.

- Bagozzi, R. P. (1993). An examination of the psychometric properties of measures of negative affect in the PANAS-X scales. *Journal of Personality and Social Psychology*, 65(4), 836–851.
- Baldwin, J., & DeSouza, E. (2001). Modelo de Maria and machismo: The social construction of gender in Brazil. *Revista Interamericana de Psicologia*, 35(1), 9–29.
- Banerjee, K., Haque, O. S., & Spelke, E. S. (2013). Melting lizards and crying mailboxes: children's preferential recall of minimally counterintuitive concepts. *Cognitive Science*, 37(7), 1251–1289.
- Banerjee, P., Chatterjee, P., & Sinha, J. (2012). Is it light or dark? Recalling moral behavior changes perception of brightness. *Psychological Science*, 23(4), 407–9.
- Baron-Cohen, S. (1999). Evolution of a theory of mind ? In M. Corballis & S. Lea (Eds.) *The Descent of the Mind: Psychological Perspectives on Hominid Evolution*. (pp.31-45) Oxford, UK: Oxford University Press.
- Baron-Cohen, S. (2000). Theory of Mind and Autism : A Review. *International Review of Research in Mental Retardation*, 23, 169–184.
- Barrett, J. L. (1998). Cognitive Constraints on Hindu Concepts of the Divine. *Journal for the Scientific Study of Religion*, 37(4), 608–619.
- Barrett, J. L. (1999). Theological Correctness: Cognitive Constraint and the Study of Religion. *Method Theory in the Study of Religion*, 11(4), 325–339.
- Barrett, J. L. (2004). *Why would anyone believe in God? Cognitive science of religion series*. Walnut Creek, CA: AltaMira Press.
- Barrett, J. L. (2005). In the Empirical Mode: Evidence needed for the Modes of Religiosity. In H. Whitehouse & R. N. McCauley (Eds.), *Mind and Religion: Psychological and Cognitive Foundations of Religiosity*. Walnut Creek, CA: AltaMira Press.
- Barrett, J. L. (2007). Cognitive science of religion: What is it and why is it? *Religion Compass*, 1(6), 768–786.
- Barrett, J. L. (2008). Coding and Quantifying Counterintuitiveness in Religious Concepts: Theoretical and Methodological Reflections. *Method Theory in the Study of Religion*, 20(4), 308–338.
- Barrett, J. L. (2011). Cognitive science of religion : Looking back. *Journal for the Scientific Study of Religion*, 50(2), 229–239.
- Barrett, J. L. (2012). *Born Believers: The Science of Children's Religious Belief*. New York, NY: Free Press.
- Barrett, J. L., & Keil, F. C. (1996). Conceptualizing a nonnatural entity: Anthropomorphism in God concepts. *Cognitive Psychology*, 31(3), 219–47.
- Barrett, J. L., & Lawson, E. T. (2001). Ritual Intuitions: Cognitive Contributions to Judgments of Ritual Efficacy. *Journal of Cognition and Culture*, 1(2), 183–201.
- Barrett, L. F. (2006). Are emotions natural kinds? *Perspectives on Psychological Science*, 1(1), 28–58.
- Barrett, L. F., Lindquist, K. A., Bliss-Moreau, E., Duncan, S., Gendron, M., Mize, J., & Brennan, L. (2007). Of mice and men: Natural kinds of emotions in the mammalian brain? A response to Panksepp and Izard. *Perspectives on Psychological Science*, 2(3), 297–312.
- Barsade, S. G. (2002). The ripple effect: Emotional contagion and its influence on group behavior. *Administrative Science Quarterly*, 47(4), 644–675.
- Barth, F. (2002). Review of Arguments and Icons. *Journal of Ritual Studies*, 16, 14–17.
- Bastian, B., Jetten, J., & Fasoli, F. (2011). Cleansing the soul by hurting the flesh: the guilt-reducing effect of pain. *Psychological Science*, 22(3), 334–5.
- Bastian, B., Jetten, J., & Ferris, L. J. (2014). Pain as social glue: shared pain increases cooperation. *Psychological Science*, 25(11), 2079–85.
- Bastian, B., Jetten, J., Hornsey, M. J., & Leknes, S. (2014). The positive consequences of pain: A biopsychosocial approach. *Personality and Social Psychology Review*, 18(3), 256–279.

- Batara, J. B. L. (2015). Religious song as a facilitator of prosocial behavior. *International Journal of Research Studies in Psychology*, 5(1).
- Bates, J. E., Freeland, C. A. B., & Lounsbury, M. L. (1979). Measurement of infant difficultness. *Child Development*, 794–803.
- Batson, C. D., Fultz, J., & Schoenrade, P. A. (1987). Distress and empathy: Two qualitatively distinct vicarious emotions with different motivational consequences. *Journal of Personality*, 55(1), 19–39.
- Batson, C. D., O'Quin, K., Fultz, J., Vanderplas, M., & Isen, A. M. (1983). Influence of self-reported distress and empathy on egoistic versus altruistic motivation to help. *Journal of Personality and Social Psychology*, 45(3), 706–718.
- Baumard, N., & Boyer, P. (2013). Religious Beliefs as Reflective Elaborations on Intuitions: A Modified Dual-Process Model. *Current Directions in Psychological Science*, 22(4), 295–300.
- Baumard, N., Hyafil, A., Morris, I., & Boyer, P. (2014). Increased Affluence Explains the Emergence of Ascetic Wisdoms and Moralizing Religions. *Current Biology*, 25(1), 10–15.
- Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, 5(4), 323–370.
- Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, 117(3), 497–529.
- Beatty, A. (2005). Emotions in the field: What are we talking about? *Journal of the Royal Anthropological Institute*, 11(1), 17–37.
- Beatty, A. (2014). Anthropology and emotion. *Journal of the Royal Anthropological Institute*, 20(3), 545–563.
- Beatty, S. E., Yoon, M.-H., Grunert, S. C., & Helgeson, J. G. (1996). An examination of gift-giving behaviors and personal values in four countries. *Gift-Giving: A Research Anthology*, 19–36.
- Becker, D. V., Anderson, U. S., Mortensen, C. R., Neufeld, S. L., & Neel, R. (2011). The face in the crowd effect unconfounded: happy faces, not angry faces, are more efficiently detected in single-and multiple-target visual search tasks. *Journal of Experimental Psychology: General*, 140(4), 637.
- Befu, H. (1968). Gift-giving in a modernizing Japan. *Monumenta Nipponica*, 445–456.
- Befu, H. (1980). A critique of the group model of Japanese society. *Social Analysis*, (5/6), 29–43.
- Befu, H. (1992). Symbols of nationalism and nihonjinron. *Ideology and Practice in Modern Japan*, 26–46.
- Befu, H. (2001). *Hegemony of homogeneity: an anthropological analysis of "Nihonjinron"* (Vol. 5). Melbourne: Trans Pacific Press.
- Bell, C. M. (1992). *Ritual Theory, Ritual Practice*. New York: Oxford University Press.
- Ben-Ari, E., Moeran, B., & Valentine, J. (1990). *Unwrapping Japan: Society and culture in anthropological perspective*. Manchester, UK: Manchester University Press.
- Benedict, R. (1934). *Patterns of Culture*. Boston, NY: Houghton Mifflin Company.
- Benedict, R. (2006). *The Chrysanthemum and the Sword*. Boston, NY: Mariner Books.
- Bering, J. M. (2002). Intuitive Conceptions of Dead Agents' Minds: The Natural Foundations of Afterlife Beliefs as Phenomenological Boundary. *Journal of Cognition and Culture*, 2(4), 263–308.
- Bering, J. M. (2010). *The God Instinct: The Psychology of Souls, Destiny and the Meaning of Life*. Boston, MA: Nicholas Brealey Publishing.
- Bering, J. M., & Bjorklund, D. F. (2004). The natural emergence of reasoning about the afterlife as a developmental regularity. *Developmental psychology*, 40(2), 217.
- Bering, J. M., McLeod, K., & Shackelford, T. K. (2005). Reasoning about dead agents reveals possible adaptive trends. *Human Naturean Interdisciplinary Biosocial Perspective*, 16(4), 360–381.

- Bernieri, F. J., & Rosenthal, R. (1991). Interpersonal coordination: Behavior matching and interactional synchrony. In R.S. Feldman & B. Rime (Eds.) *Fundamentals of nonverbal behavior: Studies in emotion & social interaction*. (pp.401-432), New York, NY: Cambridge University Press.
- Billig, M. G., & Tajfel, H. (1973). Social categorization and similarity in intergroup behaviour. *European Journal of Social Psychology*, 3(1), 27–52.
- BJJ Heroes. (2015). The Gauntlet (Belt Whipping). *BJJ Heroes.com*. Available at: <http://www.bjjheroes.com/bjj-culture/the-gauntlet-belt-whipping> [Last Accessed 3/3/2016]
- Bloch, M. (2014, June). Why there should not be a cognitive theory of “religion.” Presentation at *The 5th Conference and General Assembly of the International Association for the Cognitive Science of Religion*. Brno, Czech Republic.
- Bloom, P. (2004). *Descartes’ baby: How the Science of Child Development Explains What Makes Us Human*. New York, NY: Basic Books.
- Blount, B. G. (2011). A history of cognitive anthropology. In D. B. Kronenfeld, G. Bennardo, V. C. de Munck, & M. D. Fischer (Eds.), *A Companion to Cognitive Anthropology*. Oxford, UK: Wiley-Blackwell Publishing Ltd.
- Boecker, H., Sprenger, T., Spilker, M. E., Henriksen, G., Koppenhoefer, M., Wagner, K. J., ... Tolle, T. R. (2008). The runner’s high: opioidergic mechanisms in the human brain. *Cerebral Cortex*, 18(11), 2523–2531.
- Bohannon, J. N. (1988). Flashbulb memories for the space shuttle disaster: A tale of two theories. *Cognition*, 29(2), 179–196.
- Borisuit, A., Linhart, F., Scartezzini, J.-L., & Munch, M. (2014). Effects of realistic office daylighting and electric lighting conditions on visual comfort, alertness and mood. *Lighting Research and Technology*, 47(2), 192–209
- Bornstein, G., Crum, L., Wittenbraker, J., Harring, K., Insko, C. A., & Thibaut, J. (1983). On the measurement of social orientations in the minimal group paradigm. *European Journal of Social Psychology*, 13(4), 321–350.
- Borofsky, R. (1997). Cook, Lono, Obeyesekere, and Sahlins. *Current Anthropology*, 38(2), 255–282.
- Boukydis, C. F. Z., & Burgess, R. L. (1982). Adult Physiological Response to Infant Cries: Effects of Temperament of Infant, Parental Status, and Gender. *Child Development*, 53(5), 1291–1298.
- Bourhis, R. Y., Sachdev, I., & Gagnon, A. (1994). Intergroup research with the Tajfel matrices: Methodological notes. In *The psychology of prejudice: The Ontario symposium Vol. 7*. (pp. 209–232). Mahwah, NJ: Lawrence Erlbaum.
- Boyd, R., & Richerson, P. J. (1976). Simple Dual Inheritance Model of Conflict between Social and Biological Evolution. *Zygon*, 11(3), 254–262.
- Boyd, R., & Richerson, P. J. (1985). *Culture and the Evolutionary Process*. Chicago, IL: University of Chicago Press.
- Boyer, P. (1994). *The Naturalness of Religious Ideas: A Cognitive Theory of Religion*. Berkeley, CA: University of California Press.
- Boyer, P. (1996). What makes anthropomorphism natural: Intuitive ontology. *Journal of the Royal Anthropological Institute*, 2(1), 83.
- Boyer, P. (2000). Evolution of the modern mind and the origins of culture: religious concepts as a limiting-case. *Evolution and the Human Mind Modularity Language and Metacognition*, 93.
- Boyer, P. (2001). *Religion Explained: The Evolutionary Origins of Religious Thought*. Religion. New York, NY: Basic Books.
- Boyer, P. (2011). Why would (otherwise intelligent) scholars believe in “Religion”? *International cognition and culture institute*. Available at: <http://www.cognitionandculture.net/home/blog/35-pascals-blog/764-why-would-otherwise-intelligent-scholars-believe-in-qreligionq> [Last Accessed 3/3/2016]

- Boyer, P. (2013). Explaining religious concepts. Lévi-Strauss the brilliant and problematic ancestor. In D. Xygalatas & L. McCorkle (Eds.), *Mental Culture, Classical Social Theory and the Cognitive Science of Religion*. (pp. 164–75). Durham, UK: Acumen.
- Boyer, P., & Liénard, P. (2006). Why ritualized behavior? Precaution systems and action parsing in developmental, pathological and cultural rituals. *Behavioral and Brain Sciences*, 29(06), 595–
- Boyer, P., & Ramble, C. (2001). Cognitive templates for religious concepts: cross-cultural evidence for recall of counter-intuitive representations. *Cognitive Science*, 25(4), 535–564.
- Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: the Self-Assessment Manikin and the Semantic Differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25(1), 49–59.
- Bradley, M. M., & Lang, P. J. (1999). *The International affective digitized sounds (IADS): stimuli, instruction manual and affective ratings*. NIMH Center for the Study of Emotion and Attention.
- Brand, A. G. (1985). Hot cognition: Emotions and writing behavior. *Journal of Advanced Composition*, 5–15.
- Bressan, P., & Kramer, P. (2015). Human kin detection. *Wiley Interdisciplinary Reviews: Cognitive Science*, 6(3), 299–311.
- Brewer, M. B. (1991). The Social Self: On Being the Same and Different at the Same Time. *Personality and Social Psychology Bulletin*, 17(5), 475–482.
- Brewer, M. B., & Gardner, W. (1996). Who is this “We”? Levels of collective identity and self representations. *Journal of Personality and Social Psychology*, 71(1), 83–93.
- Brown, J. L., Sheffield, D., Leary, M. R., & Robinson, M. E. (2003). Social support and experimental pain. *Psychosomatic Medicine*, 65(2), 276–83.
- Brown, R., & Kulik, J. (1977). Flashbulb memories. *Cognition*, 5(1), 73–99.
- Bruce, V., & Young, A. (1998). *In the eye of the beholder: the science of face perception*. Oxford, UK: Oxford University Press.
- Brugger, A., Lariviere, L. A., Mumme, D. L., & Bushnell, E. W. (2007). Doing the right thing: Infants’ selection of actions to imitate from observed event sequences. *Child Development*, 78(3), 806–824.
- Buhrmester, M. D. (2013, August 1). *Understanding the cognitive and affective underpinnings of whistleblowing*. University of Texas. [Doctoral Thesis]
- Buhrmester, M. D., Fraser, W. T., Lanman, J. A., Whitehouse, H., & Swann, W. B. (2014). When terror hits home: Fused Americans saw Boston bombing victims as “family” and rushed to their aid. *Self and Identity*, 14(3), 253–270.
- Buhrmester, M. D., Gómez, Á., Brooks, M. L., Morales, J. F., Fernández, S., & Swann, W. B. (2012). My Group’s Fate Is My Fate: Identity-Fused Americans and Spaniards Link Personal Life Quality to Outcome of 2008 Elections. *Basic and Applied Social Psychology*, 34(6), 527–533.
- Bulbulia, J. A. (2004a). Religious Costs as Adaptations. *Evolution and Cognition*, 10(1), 19–42.
- Bulbulia, J. A. (2004b). The cognitive and evolutionary psychology of religion. *Biology and Philosophy*, 19(5), 655–686.
- Bulbulia, J. A. (2005). Are There Any Religions? An Evolutionary Exploration. *Method & Theory in the Study of Religion*, 17(2), 71–100.
- Bulbulia, J. A. (2007). The evolution of religion. In R. I. M. Dunbar & L. Barrett (Eds.), *Oxford handbook of evolutionary psychology* (Vol. 21, pp. 621–636). Oxford, UK: Oxford University Press.
- Bulbulia, J. A. (2008). Meme Infection or Religious Niche Construction? An Adaptationist Alternative to The Cultural Maladaptationist Hypothesis. *Method & Theory in the Study of Religion*, 20(1), 67–107.
- Bulbulia, J. A., Geertz, A. W., Atkinson, Q. D., Cohen, E., Evan, N., Francois, P., ... Wilson, D. S. (2013). The Cultural Evolution of Religion. In P. J. Richerson & M. H. Christiansen (Eds.), *Cultural Evolution: Science, Technology, Language, and Religion*. (pp. 381–404). Cambridge, MA: MIT Press.

- Bulbulia, J. A., & Sosis, R. (2011). Signalling theory and the evolution of religious cooperation. *Religion, 41*(3), 363–388.
- Cacioppo, J. T., Berntson, G. G., Larsen, J. T., Poehlmann, K. M., & Ito, T. A. (2000). The psychophysiology of emotion. *Handbook of Emotions, 2*, 173–191.
- Cacioppo, J. T., Petty, R. E., Losch, M. E., & Kim, H. S. (1986). Electromyographic activity over facial muscle regions can differentiate the valence and intensity of affective reactions. *Journal of Personality and Social Psychology, 50*(2), 260–8.
- Campo, S., Poulos, G., & Sipple, J. W. (2005). Prevalence and profiling: Hazing among college students and points of intervention. *American Journal of Health Behavior, 29*(2), 137–149.
- Cardoso, C., Ellenbogen, M. A., Serravalle, L., & Linnen, A.-M. (2013). Stress-induced negative mood moderates the relation between oxytocin administration and trust: evidence for the tend-and-befriend response to stress? *Psychoneuroendocrinology, 38*(11), 2800–2804.
- Carifio, J., & Perla, R. J. (2007). Ten Common Misunderstandings, Misconceptions, Persistent Myths and Urban Legends about Likert Scales and Likert Response Formats and their Antidotes. *Journal of Social Sciences, 3*(3), 106–116.
- Carver, C. S., & Scheier, M. F. (1990). Origins and functions of positive and negative affect: a control-process view. *Psychological Review, 97*(1), 19.
- Casler, K., & Kelemen, D. (2008). Developmental Continuity in Teleo-Functional Explanation: Reasoning about Nature Among Romanian Romani Adults. *Journal of Cognition and Development, 9*(3), 340–362.
- Cason, T. N., & Khan, F. U. (1999). A laboratory study of voluntary public goods provision with imperfect monitoring and communication. *Journal of Development Economics, 58*(2), 533–552.
- Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research, 1*(2), 245–276.
- Chapman, K. P. (2005). *Inside the Dojo: Participation and Performance in the Japanese Martial Arts*. School of Oriental & African Studies.[Doctoral Thesis]
- Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception–behavior link and social interaction. *Journal of Personality and Social Psychology, 76*(6), 893.
- Chen, D. L. (2010). Club Goods and Group Identity: Evidence from Islamic Resurgence during the Indonesian Financial Crisis. *Journal of Political Economy, 118*(2), 300–330.
- Cheng, G. H. L., Fielding, K. S., & Terry, D. J. (2011). Responses of group members to procedural discrimination: The role of ingroup prototypicality. *Group Processes Intergroup Relations, 14*(4), 461–476.
- Chiou, J.-S. (2001). Horizontal and vertical individualism and collectivism among college students in the United States, Taiwan, and Argentina. *The Journal of Social Psychology, 141*(5), 667–678.
- Chiou, W.B., & Cheng, Y.Y. (2013). In broad daylight, we trust in God! Brightness, the salience of morality, and ethical behavior. *Journal of Environmental Psychology, 36*, 37–42.
- Cho, F., & Squier, R. K. (2008a). He Blinded Me with Science: Science Chauvinism in the Study of Religion. *Journal of the American Academy of Religion, 76*(2), 420–448.
- Cho, F., & Squier, R. K. (2008b). Reductionism: Be Afraid, Be Very Afraid. *Journal of the American Academy of Religion, 76*(2), 412–417.
- Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Christianson, S. A. (1989). Flashbulb memories: special, but not so special. *Memory & Cognition, 17*, 435–443.
- Cimino, A. (2011). The evolution of hazing: Motivational mechanisms and the abuse of newcomers. *Journal of Cognition and Culture, 11*(3-4), 241–267.
- Cimino, A. (2013). Predictors of hazing motivation in a representative sample of the United States. *Evolution and Human Behavior, 34*(6), 446–452.

- Cimino, A., & Delton, A. W. (2010). On the Perception of Newcomers. *Human Nature*, 21(2), 186–202.
- Claire, L. St., & Turner, J. C. (1982). The role of demand characteristics in the social categorization paradigm. *European Journal of Social Psychology*, 12(3), 307–314.
- Clark, L. A., & Watson, D. (1989). *The Japanese Positive and Negative Affect Schedule: Factor-based scales for the assessment of mood*. Available at: <http://www2.psychology.uiowa.edu/faculty/clark/J-PANAS.pdf> [Last Accessed 3/3/2016]
- Clegg, J. M., & Legare, C. H. (2015). Instrumental and conventional interpretations of behavior are associated with distinct outcomes in early childhood. *Child Development*, 87(2), 527–542.
- Clifford, J., & Marcus, G. E. (1986). *Writing Culture*. (J. Clifford & G. Marcus, Eds.) *Culture* (Vol. 5). University of California Press.
- Cogan, R., & Spinnato, J. A. (1986). Pain and discomfort thresholds in late pregnancy. *Pain*, 27(1), 63–8.
- Cohen, E. (2007). *The Mind Possessed: The cognition of spirit possession in an Afro-Brazilian religious tradition*. New York, NY: Oxford University Press.
- Cohen, E., Ejsmond-Frey, R., Knight, N., & Dunbar, R. I. M. (2010). Rowers' high: behavioural synchrony is correlated with elevated pain thresholds. *Biology Letters*, 6(1), 106–8.
- Cohen, E., Lanman, J. A., Whitehouse, H., & McCauley, R. N. (2008). Common criticisms of the cognitive science of religion - answered. *Bulletin of the Council of Societies for the Study of Religion*, 37(4), 112–115.
- Collins, P. (2005). Thirteen Ways of Looking at a Ritual. *Journal of Contemporary Religion*, 20(3), 323–342.
- Collins, P. (2009). Religion and ritual: A multi-perspective approach. In P. B. Clarke (Ed.), *The Oxford handbook of the sociology of religion*. (pp. 671–687). Oxford, UK: Oxford University Press.
- Conklin, H. C. (1962). Lexicographical Treatment of Folk Taxonomies. In F. W. Householder & S. Saporta (Eds.), *Problems in Lexicography*. Bloomington: Indiana Research Center in Anthropology, Folklore, and Linguistics.
- Conway, M. A. (1995). *Flashbulb Memories: Essays in Cognitive Psychology*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Correia, D. (2013). F**k Jared Diamond. *Capitalism Nature Socialism*, 24(4), 1–6.
- Covell, S. G. (2005). *Japanese Temple Buddhism: Worldliness in a Religion of Renunciation*. Honolulu, HI: University of Hawaii Press.
- Cozma, I. (2011). How are individualism and collectivism measured. *Romanian Journal of Applied Psychology*, 13(1), 11–17.
- Crawford, J. R., & Henry, J. D. (2004). The positive and negative affect schedule (PANAS): construct validity, measurement properties and normative data in a large non-clinical sample. *The British Journal of Clinical Psychology*, 43(3), 245–265.
- Croson, R. T. A., & Marks, M. B. (2000). Step returns in threshold public goods: A meta-and experimental analysis. *Experimental Economics*, 2(3), 239–259.
- Csordas, T. J. (1990). Embodiment as a Paradigm for Anthropology. *Ethos*, 18(1), 5–47.
- Cubelli, R., & Della Sala, S. (2008). Flashbulb memories: special but not iconic. *Cortex*, 44(7), 908–
- Curci, A., & Conway, M. A. (2013). Playing the flashbulb memory game: a comment on Cubelli and Della Sala. *Cortex*, 49(1), 356–7.
- Czachesz, I. (2010). Long term, explicit memory in rituals. *Journal of Cognition and Culture*, 10(3–4), 327–339.
- Damasio, A. (2008). *Descartes' Error: Emotion, Reason and the Human Brain*. New York, NY: Random House.
- Damisch, L., Stoberock, B., & Mussweiler, T. (2010). Keep your fingers crossed!: how superstition improves performance. *Psychological Science*, 21(7), 1014–20.

- Darley, J. M., & Aronson, E. (1966). Self-evaluation vs. direct anxiety reduction as determinants of the fear-affiliation relationship. *Journal of Experimental Social Psychology*, 1, 66–79.
- Darwin, C. (1965). *The expression of the emotions in man and animals* (Vol. 526). Chicago, IL: University of Chicago press.
- Davidson, P. S. R., & Glisky, E. L. (2002). Is flashbulb memory a special instance of source memory? Evidence from older adults. *Memory*, 10(2), 99–111.
- de Albuquerque, C. L., & Paes-Machado, E. (2004). The hazing machine: the shaping of Brazilian military police recruits. *Policing and Society*, 14(2), 175–192.
- De Waal, F. B. M. (1996). *Good natured*. New York, NY: Harvard University Press.
- de Winter, J. C. F. (2013). Using the Student's t-test with extremely small sample sizes. *Practical Assessment, Research & Evaluation*, 18(10), 1–12.
- DeBruine, L. M. (2005). Trustworthy but not lust-worthy: context-specific effects of facial resemblance. *Proceedings of the Royal Society of London B: Biological Sciences*, 272(1566), 317–335.
- Delton, A. W., & Cimino, A. (2010). Exploring the evolved concept of NEWCOMER: Experimental tests of a cognitive model. *Evolutionary Psychology*, 8(2), 317–335.
- Derrida, J. (1997). *Of Grammatology*. Baltimore, NY: The John Hopkins University Press.
- Dezecache, G., & Dunbar, R. I. M. (2012). Sharing the joke: the size of natural laughter groups. *Evolution and Human Behavior*, 33(6), 775–779.
- Dickerson, S. S., & Kemeny, M. E. (2004). Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. *Psychological Bulletin*, 130(3), 355–391.
- Diehl, M. (1990). The Minimal Group Paradigm: Theoretical Explanations and Empirical Findings. *European Review of Social Psychology*, 1(1), 263–292.
- Dinneen, L. C., & Blakesley, B. C. (1973). Algorithm AS 62: A generator for the sampling distribution of the Mann-Whitney U statistic. *Applied Statistics*, 269–273.
- Dore, R. P. (1973). *British Factory, Japanese Factory: The Origins of National Diversity in Industrial Relations*. London, UK: Allen and Unwin.
- Douglas, M. (1966). *Purity and Danger: An analysis of concepts of pollution and taboo*. London, UK: Routledge.
- Dow, J. (2008). Is Religion an Evolutionary Adaptation? *Journal of artificial societies and social simulation*, 11(2), 2.
- Dunbar, R. I. M. (2003). The Social Brain: Mind, Language, and Society in Evolutionary Perspective. *Annual Review of Anthropology*, 32(1), 163–181.
- Dunbar, R. I. M. (2009). The social brain hypothesis and its implications for social evolution. *Annals of Human Biology*, 36(5), 562–72.
- Dunn, L., & Hoegg, J. (2014). The Impact of Fear on Emotional Brand Attachment. *Journal of Consumer Research*, 41(1), 152–168.
- Durkheim, E. (1965). *The Elementary forms of the religious life*. [J. M. Ward, Trans.]. London, UK: Free Press.
- Earhart, D. (2005). All Ready To Die: Kamekazefication and Japan's Wartime Ideology. *Critical Asian Studies*, 37(4), 569–596.
- Edwards, L., & Torcellini, P. A. (2002). *A literature review of the effects of natural light on building occupants*. Golden, CO: National Renewable Energy Laboratory.
- Eisenberger, N. I., & Lieberman, M. D. (2004). Why rejection hurts: a common neural alarm system for physical and social pain. *Trends in Cognitive Sciences*, 8(7), 294–300.
- Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion*, 6(3-4), 169–200.
- Ekman, P. (1994). Strong evidence for universals in facial expressions: a reply to Russell's mistaken critique. *Psychological Bulletin*, 115(2), 268–287.
- Ekman, P., & Cordaro, D. (2011). What is meant by calling emotions basic. *Emotion Review*, 3(4), 364–370.

- Ekman, P., Friesen, W. V., O'Sullivan, M., Chan, A., Diacoyanni-Tarlatzis, I., Heider, K. G., ... Ricci-Bitti, P. E. (1987). Universals and cultural differences in the judgments of facial expressions of emotion. *Journal of Personality and Social Psychology*, 53(4), 712-717.
- Ekman, P., & Heider, K. G. (1988). The universality of a contempt expression: A replication. *Motivation and Emotion*, 12(3), 303-308.
- Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: a meta-analysis. *Psychological Bulletin*, 128(2), 203-235.
- Er, N. (2003). A new flashbulb memory model applied to the Marmara earthquake. *Applied Cognitive Psychology*, 17(5), 503-517.
- Evans, J. S. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, 59, 255-278.
- Falk, C. F., Heine, S. J., & Takemura, K. (2013). Cultural Variation in the Minimal Group Effect. *Journal of Cross-Cultural Psychology*, 45(2), 265-281.
- Festinger, L. (1962). *A theory of cognitive dissonance*. Redwood City, CA: Stanford University Press.
- Finkel, M. A. (2002). Traumatic injuries caused by hazing practices. *The American Journal of Emergency Medicine*, 20(3), 228-233.
- Fischer, R., Callander, R., Reddish, P., & Bulbulia, J. A. (2013). How do rituals affect cooperation? An experimental field study comparing nine ritual types. *Human Nature*, 24(2), 115-25.
- Fischer, R., & Xygalatas, D. (2014). Extreme Rituals as Social Technologies. *Journal of Cognition and Culture*, 14(5), 345-355.
- Fischer, R., Xygalatas, D., Mitkidis, P., Reddish, P., Tok, P., Konvalinka, I., & Bulbulia, J. A. (2014). The fire-walker's high: affect and physiological responses in an extreme collective ritual. *PLoS One*, 9(2), e88355.
- Fitzgerald, T. (1999). *The Ideology of Religious Studies*. Oxford, UK: Oxford University Press.
- Fitzgerald, T. (2003). Religion and the Secular in Japan: Problems in history, social anthropology and the study of religion. *Electronic journal of contemporary japanese studies*, 1.
- Fitzgerald, T. (2004). The Religion-Secular Dichotomy: A Response to Responses. *Electronic journal of contemporary japanese studies*, 4.
- Fleming, A. S., Corter, C., Stallings, J., & Steiner, M. (2002). Testosterone and Prolactin Are Associated with Emotional Responses to Infant Cries in New Fathers. *Hormones and Behavior*, 42(4), 399-413.
- Fleming, A. S., Steiner, M., & Corter, C. (1997). Cortisol, hedonics, and maternal responsiveness in human mothers. *Hormones and Behavior*, 32(2), 85-98.
- Frake, C. O. (1969). The Ethnographic Study of Cognitive Systems. In S. Tyler (ed.) *Cognitive Anthropology*, (pp. 28-41) New York, NY: Holt, Rinehart, & Winston.
- Franek, J. (2015). Has the cognitive science of religion (re)defined "religion"? *Religio*, 22(1), 3-27.
- Frank, M. G., & Gilovich, T. (1988). The dark side of self-and social perception: black uniforms and aggression in professional sports. *Journal of Personality and Social Psychology*, 54(1), 74.
- Frazer, J. G. (1922). *The Golden Bough: Volumes I-XII*. New York, NY: The Macmillan Co.
- Freeman, D. (1983). *Margaret Mead and Samoa: The making and unmaking of an anthropological myth*. Cambridge, MA: Harvard University Press.
- Frith, C., & Frith, U. (2005). Theory of mind. *Current Biology*, 15(17), R644-R645.
- Fukuyama, F. (2011). *The Origins of Political Order: From Prehuman Times to the French Revolution*. New York, NY: Macmillan.
- Gaertner, L., & Insko, C. A. (2001). On the measurement of social orientations in the minimal group paradigm: norms as moderators of the expression of intergroup bias. *European Journal of Social Psychology*, 31(2), 143-154.
- Gagnon, A., & Bourhis, R. Y. (1996). Discrimination in the minimal group paradigm: Social identity or self-interest? *Personality and Social Psychology Bulletin*, 22(12), 1289-1301.

- Galasiu, A. D., & Veitch, J. A. (2006). Occupant preferences and satisfaction with the luminous environment and control systems in daylit offices: a literature review. *Energy and Buildings*, 38(7), 728–742.
- Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, 2(12), 493–501.
- Garcia, H. (2011). *A Geek in Japan: Discovering the Land of Manga, Anime, Zen, and the Tea Ceremony*. Clarendon, VM: Tuttle Publishing.
- Geertz, C. (1973). *The Interpretation of Cultures*. New York, NY: Basic Books.
- Gerard, H. B., & Mathewson, G. C. (1966). The effects of severity of initiation on liking for a group: A replication. *Journal of Experimental Social Psychology*, 2(3), 278–287.
- Gergely, G., Bekkering, H., & Király, I. (2002). Rational imitation in preverbal infants. *Nature*, 415(6873), 755.
- Gergely, G., & Csibra, G. (2005). The social construction of the cultural mind: Imitative learning as a mechanism of human pedagogy. *Interaction Studies*, 6(3), 463–481.
- Gervais, W. M., Willard, A. K., Norenzayan, A., & Henrich, J. (2011). The Cultural Transmission of Faith: Why innate intuitions are necessary, but insufficient, to explain religious belief. *Religion*, 41(3), 389–410.
- Giardino, J., Gonzalez, A., Steiner, M., & Fleming, A. S. (2008). Effects of motherhood on physiological and subjective responses to infant cries in teenage mothers: a comparison with non-mothers and adult mothers. *Hormones and Behavior*, 53(1), 149–158.
- Ginges, J., Hansen, I., & Norenzayan, A. (2009). Religion and support for suicide attacks. *Psychological Science*, 20(2), 224–30.
- Givon, T. (1995). *Functionalism and Grammar*. Philadelphia, PA: Benjamins.
- Glass, G. V., Peckham, P. D., & Sanders, J. R. (1972). Consequences of Failure to Meet Assumptions Underlying the Fixed Effects Analyses of Variance and Covariance. *Review of Educational Research*, 42(3), 237–288.
- Gomes, C. M., & McCullough, M. E. (2015). The effects of implicit religious primes on dictator game allocations: A preregistered replication experiment. *Journal of Experimental Psychology: General*, 144(6), e94.
- Gómez, A., Brooks, M. L., Buhrmester, M. D., Vázquez, A., Jetten, J., & Swann, W. B. (2011). On the nature of identity fusion: insights into the construct and a new measure. *Journal of Personality and Social Psychology*, 100(5), 918–33.
- Gómez, A., Morales, J. F., Hart, S., Vázquez, A., & Swann, W. B. (2011). Rejected and excluded forevermore, but even more devoted: irrevocable ostracism intensifies loyalty to the group among identity-fused persons. *Personality & Social Psychology Bulletin*, 37(12), 1574–86.
- Gou, Z., Lau, S. S.-Y., & Qian, F. (2013). Comparison of mood and task performance in naturally-lit and artificially-lit environments. *Indoor and Built Environment*, 24(1), 27–36.
- Gouveia, V. V., Clemente, M., & Espinosa, P. (2003). The horizontal and vertical attributes of individualism and collectivism in a Spanish population. *The Journal of Social Psychology*, 143(1), 43–63.
- Graham, S. M., Huang, J. Y., Clark, M. S., & Helgeson, V. S. (2008). The positives of negative emotions: willingness to express negative emotions promotes relationships. *Personality & Social Psychology Bulletin*, 34(3), 394–406.
- Grammer, K., & Eibl-Eibesfeldt, I. (1990). The ritualisation of laughter. *Natürlichkeit Der Sprache Und Der Kultur*, 18, 192–214.
- Green, D. P., Goldman, S. L., & Salovey, P. (1993). Measurement error masks bipolarity in affect ratings. *Journal of Personality and Social Psychology*, 64(6), 1029–41.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: the implicit association test. *Journal of Personality and Social Psychology*, 74(6), 1464–80.

- Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, 85(2), 197–216.
- Gregory, J. P. (submitted). Gods and talking animals: The pan-cultural transmission advantage of supernatural agent concepts. *Journal of Cognition and Culture*.
- Gregory, J. P. (forthcoming). Is there a window of opportunity in childhood and adolescence for religiosity? Children and adolescents of UK and China preferentially recall religious-type cultural representations, but older adults do not. *Religion, Brain & Behavior*.
- Gregory, J. P., & Barrett, J. L. (2009). Epistemology and Counterintuitiveness: Role and Relationship in Epidemiology of Cultural Representations. *Journal of Cognition and Culture*, 9(3), 289–314.
- Grillon, C. (2008). Models and mechanisms of anxiety: evidence from startle studies. *Psychopharmacology*, 199(3), 421–437.
- Grillon, C., Pellowski, M., Merikangas, K. R., & Davis, M. (1997). Darkness facilitates the acoustic startle reflex in humans. *Biological Psychiatry*, 42(6), 453–460.
- Grimes, R. L. (2013). *The Craft of Ritual Studies*. Oxford University Press.
- Groves, M., Griggs, G., & Leflay, K. (2011). Hazing and initiation ceremonies in university sport: setting the scene for further research in the United Kingdom. *Sport in Society*, 15(1), 117–131.
- Guo, X., Schwartz, S. J., & McCabe, B. E. (2008). Aging, gender, and self: Dimensionality and measurement invariance analysis on self-construal. *Self and Identity*, 7(1), 1–24.
- Guthrie, S. E. (1980). A Cognitive Theory of Religion. *Current Anthropology*, 21(2), 181–203.
- Guthrie, S. E. (1993). *Faces in the Clouds: A New Theory of Religion*. New York, NY: Oxford University Press.
- Guthrie, S. E. (2013). Early cognitive theorists of religion: Robert Horton and his predecessors. In D. Xygalatas & W. W. McCorkle (Eds.), *Mental Culture: Classical Social Theory and the Cognitive Science of Religion*. (pp. 33–51). London, UK: Acumen.
- Haidt, J. (2001). The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychological Review*, 108(4), 814–834.
- Haidt, J., Patrick Seder, J., & Kesebir, S. (2008). Hive Psychology, Happiness, and Public Policy. *The Journal of Legal Studies*, 37(S2), S133–S156.
- Hardacre, H. (1986). Creating state shinto: The great promulgation campaign and the new religions. *Journal of Japanese Studies*, 29–63.
- Hardacre, H. (1988). *Kurozumikyo and the New Religions of Japan*. *Journal Of African Studies* (Vol. 47). Princeton, NJ: Princeton University Press.
- Hardacre, H. (2007). Aum Shinrikyô and the Japanese Media: The Pied Piper Meets the Lamb of God. *History of Religions*, 47(2/3), 171–204.
- Hardin, G. (1968). The tragedy of the commons. *Science*, 162(3859), 1243–1248.
- Harmon-Jones, E., & Harmon-Jones, C. (2007). Cognitive dissonance theory after 50 years of development. *Zeitschrift Für Sozialpsychologie*, 38(1), 7–16.
- Harmon-Vukić, M., Upal, M. A., & Sheehan, K. J. (2012). Understanding the memory advantage of counterintuitive concepts. *Religion, Brain & Behavior*, 2(2), 121–139.
- Harootunian, H. (1999). Memory, mourning, and national morality: Yasukuni Shrine and the reunion of state and religion in postwar Japan. In P. Van der Veer & H. Lehmann (Eds.), *Nation and Religion: Perspectives on Europe and Asia*. (pp. 144–160) Princeton, NJ: Princeton University Press.
- Harvey, G. (2013). *Food, Sex and Strangers: Understanding Religion as Everyday Life*. Durham, UK: Acumen.
- Haslam, S. A., Oakes, P. J., Turner, J. C., & McGarty, C. (1996). Social identity, self-categorization, and the perceived homogeneity of ingroups and outgroups: The interaction between social motivation and cognition. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of Motivation*

- and Cognition. *Volume 3: The Interpersonal Context*. (pp. 182–222). New York: Guildford Press.
- Hastorf, C. (2011, August 15). Steamed or Boiled: Identity and Value in Food Preparation. *eTopoi: Journal for Ancient Studies, Special Volume 2*, 213–242.
- Hatfield, E., Bensman, L., Thornton, P. D., & Rapson, R. L. (2014). New Perspectives on Emotional Contagion: A Review of Classic and Recent Research on Facial Mimicry and Contagion. *Interpersona: An International Journal on Personal Relationships*, 8(2), 159–179.
- Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1994). *Emotional contagion*. Cambridge, UK: Cambridge university press.
- Hathaway, W. E. (1995). Effects of school lighting on physical development and school performance. *The Journal of Educational Research*, 88(4), 228–242.
- Hautaluoma, J. E., & Spungin, H. (1974). Effects of Initiation Severity and Interest on Group Attitudes. *The Journal of Social Psychology*, 93(2), 245–259.
- Haxby, J. V, Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, 4(6), 223–233.
- Hayes, A. F. (2012). PROCESS: A versatile computational tool for observed variable mediation, moderation, and conditional process modeling. *White Paper*, 1–39.
- Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis*. New York, NY: Guilford.
- Haynes, B. P. (2008). The impact of office comfort on productivity. *Journal of Facilities Management*, 6(1), 37–51.
- Heine, S. J., Lehman, D. R., Markus, H. R., & Kitayama, S. (1999). Is there a universal need for positive self-regard? *Psychological Review*, 106(4), 766.
- Henrich, J. (2009). The evolution of costly displays, cooperation and religion: credibility enhancing displays and their implications for cultural evolution. *Evolution and Human Behavior*, 30(4), 244–260.
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, 33, 61–135.
- Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., ... Ziker, J. (2006). Costly punishment across human societies. *Science*, 312(5781), 1767–70.
- Hirshleifer, J. (1983). From weakest-link to best-shot: The voluntary provision of public goods. *Public Choice*, 41(3), 371–386.
- Hirst, W., Phelps, E. A., Meksins, R., Vaidya, C. J., Johnson, M. K., Mitchell, K. J., ... Olsson, A. (2015). A ten-year follow-up of a study of memory for the attack of September 11, 2001: Flashbulb memories and memories for flashbulb events. *Journal of Experimental Psychology. General*, 144(3), 604–23.
- Hogg, M. A. (2000). From prototypicality to power: A social identity analysis of leadership. *Science*, 18, 1–30.
- Hogg, M. A. (2006). Social Identity Theory. In P. J. Burke (Ed.), *Contemporary Social Psychological Theories*. (pp. 111–136). Palo Alta, CA: Stanford University Press.
- Hogg, M. A., & Turner, J. C. (1987). Intergroup behaviour, self-stereotyping and the salience of social categories. *British Journal of Social Psychology*, 26(4), 325–340.
- Holmes, R. W. (2013). *How to eradicate hazing*. Bloomington, IN: AuthorHouse.
- Hopkins, N., Reicher, S., Harrison, K., Cassidy, C., Bull, R., & Levine, M. (2007). Helping to improve the group stereotype: on the strategic dimension of prosocial behavior. *Personality & Social Psychology Bulletin*, 33(6), 776–88.
- Hornsey, M. J. (2008). Social Identity Theory and Self-categorization Theory: A Historical Review. *Social and Personality Psychology Compass*, 2(1), 204–222.
- Houseman, M. (2002). Review of Arguments and Icons. *Journal of Ritual Studies*, 16, 18–22.

- Hudenko, W. J., & Magenheimer, M. A. (2011). Listeners prefer the laughs of children with autism to those of typically developing children. *Autism*, 16(6), 641-655.
- Humphrey, C., & Laidlaw, J. (1994). *The archetypal actions of ritual*. Oxford, UK: Oxford University Press.
- Huni, G. (2014). Caio Terra Against The Gauntlet Belt Promotion Tradition in BJJ: "We Are Too Smart For That." *BJJ Eastern Europe*. Available at: <http://www.bjee.com/bjj-news/caio-terra-against-the-gauntlet-belt-promotion-tradition-in-bjj-we-are-too-smart-for-that/> [Last Accessed 3/3/2016]
- Iacoboni, M. (2009). Imitation, empathy, and mirror neurons. *Annual Review of Psychology*, 60, 653-670.
- Irons, W. (1996a). In our own self image: The evolution of morality, deception, and religion. *Skeptic*, 4(2), 50-61.
- Irons, W. (1996b). Morality as an evolved adaptation. In J. P. Hurd (Ed.), *Investigating the Biological Foundations of Morality*. (pp. 1-34). Lewiston, NY: Edwin Mellon Press.
- Irons, W. (1996c). Morality, religion and human evolution. *Religion and Science: History, Method and Dialogue*, 375-399.
- Irons, W. (2001). Religion as a hard-to-fake sign of commitment. In R. M. Nesse (Ed.), *Evolution and the Capacity for Commitment*. (pp. 292-309). New York, NY: Routledge.
- Isen, A. M. (2000). Positive Affect and Decision Making. In M. Lewis & J. Harviland-Jones (Eds.), *Handbook of emotions* (2nd Editio., pp. 417-435). New York, NY: Guildford Press.
- Isen, A. M., & Means, B. (1983). The influence of positive affect on decision-making strategy. *Social Cognition*, 2(1), 18-31.
- Isen, A. M., & Reeve, J. (2006). The Influence of Positive Affect on Intrinsic and Extrinsic Motivation: Facilitating Enjoyment of Play, Responsible Work Behavior, and Self-Control. *Motivation and Emotion*, 29(4), 295-323.
- Izard, C. E. (1994). Innate and universal facial expressions: evidence from developmental and cross-cultural research. *Psychological Bulletin*, 115(2), 288-299.
- Izard, C. E. (2007). Basic emotions, natural kinds, emotion schemas, and a new paradigm. *Perspectives on Psychological Science*, 2(3), 260-280.
- Jaffe, R. M. (2001). *Neither Monk nor Layman: Clerical Marriage in Modern Japanese Buddhism*. Princeton, NJ: Princeton University Press.
- James, W. (1884). II.—What is an emotion? *Mind*, (34), 188-205.
- Jamieson, S. (2004). Likert scales: how to (ab)use them. *Medical Education*, 38(12), 1217-8.
- Jimenez, J., Gomez, A., Buhrmester, M. D., Vazquez, A., Whitehouse, H., & Swann, W. B. (2015). The Dynamic Identity Fusion Index: A New Continuous Measure of Identity Fusion for Web-Based Questionnaires. *Social Science Computer Review*, 34, 215-228.
- Jin, N., Yamagishi, T., & Kiyonari, T. (1996). Bilateral dependency and the minimal group paradigm. *Shinrigaku Kenkyu The Japanese Journal of Psychology*, 67(2), 77-85.
- Jiu Jitsu Style. (2012 11th December). Facebook Debate: Belt whipping after promotions- hater of fan? What are the rules at your academy? Available at: <https://www.facebook.com/Meerkatsu/posts/225640460901862> [Last Accessed 3/3/2016].
- Johnson, J. (2011). Through the Liminal: A Comparative Analysis of Communitas and Rites of Passage in Sport Hazing and Initiations. *The Canadian Journal of Sociology / Cahiers Canadiens de Sociologie*, 36(3), 199-227.
- Jones, J. M., & Jetten, J. (2010). Recovering From Strain and Enduring Pain: Multiple Group Memberships Promote Resilience in the Face of Physical Challenges. *Social Psychological and Personality Science*, 2(3), 239-244.
- Jong, J., Kavanagh, C. M., & Visala, A. (2015). Born idolaters: The limits of the philosophical implications of the cognitive science of religion. *Neue Zeitschrift für Systematische Theologie und Religionsphilosophie*, 20(2), 244-266.

- Josephson, J. Å. (2011). The Invention of Japanese Religions. *Religion Compass*, 5(10), 589–597.
- Josephson, J. Å. (2012). *The Invention of Religion in Japan*. Chicago, IL: University of Chicago Press.
- Kahneman, D. (2011). *Thinking, Fast and Slow*. New York, NY: Penguin Farrar, Strauss and Giroux.
- Kaiser, H. F., & Rice, J. (1974). Little Jiffy, Mark IV. *Educational and Psychological Measurement*, 34, 111–117.
- Kamau, C. (2013). What does being initiated severely into a group do? The role of rewards. *International Journal of Psychology : Journal International de Psychologie*, 48(3), 399–406.
- Kapitány, R., & Nielsen, M. (2015). Adopting the ritual stance: The role of opacity and context in ritual and everyday actions. *Cognition*, 145, 13–29.
- Kapitány, R., & Nielsen, M. (2016). The ritual stance and the precaution system: the role of goal-demotion and opacity in ritual and everyday actions. *Religion, Brain & Behavior*, 1–16.
- Karmiloff-Smith, A. (1992). *Beyond Modularity: A Developmental Perspective on Cognitive Science*. Cambridge, MA: MIT Press.
- Karp, D., Jin, N., Yamagishi, T., & Shinotsuka, H. (1993). Raising the minimum in the minimal group paradigm. *The Japanese Journal of Experimental Social Psychology*, 32(3), 231–240.
- Kavanagh, C. M. (2013). BJJ Demographics. *BJJ Surveys.com*. Available at: <http://www.bjjsurveys.com/blog/2013/10/2/results-bjj-demographics> [Last Accessed 3/3/2016]
- Kawamura, Y., & Kare, M. R. (1987). *Umami: a basic taste. Physiology, biochemistry, nutrition, food science*. Basel, NY: Marcel Dekker, Inc.
- Keating, C. F., Pomerantz, J., Pommer, S. D., Ritt, S. J. H., Miller, L. M., & McCormick, J. (2005). Going to College and Unpacking Hazing: A Functional Approach to Decrypting Initiation Practices Among Undergraduates. *Group Dynamics: Theory, Research, and Practice*, 9(2), 104–126.
- Kelemen, D. (1999). The scope of teleological thinking in preschool children. *Cognition*, 70(3): 241–272.
- Kelemen, D. (2004). Are children “intuitive theists”? Reasoning about purpose and design in nature. *Psychological Science*, 15(5), 295–301.
- Kelemen, D., & DiYanni, C. (2005). Intuitions About Origins: Purpose and Intelligent Design in Children’s Reasoning About Nature. *Journal of Cognition and Development*, 6(1), 3–31.
- Kelemen, D., & Rosset, E. (2009). The human function compunction: teleological explanation in adults. *Cognition*, 111(1), 138–143.
- Kelly, J. R., Iannone, N. E., & McCarty, M. K. (2014). The function of shared affect in groups. In C. von Scheve & M. Salmela (Eds.), *Collective Emotions* (pp. 175–188). Oxford, UK: Oxford University Press.
- Kerr, M. (2015). *Scream: Chilling Adventures in the Science of Fear*. New York, NY: PublicAffairs.
- Keupp, S., Behne, T., & Rakoczy, H. (2013). Why do children overimitate? Normativity is crucial. *Journal of Experimental Child Psychology*, 116(2), 392–406.
- Kirkpatrick, L. A., & Shaver, P. (1988). Fear and Affiliation Reconsidered from a Stress and Coping Perspective: The Importance of Cognitive Clarity and Fear Reduction. *Journal of Social and Clinical Psychology*, 7(2-3), 214–233.
- Kline, R. B. (1998). Software review: Software programs for structural equation modeling: Amos, EQS, and LISREL. *Journal of Psychoeducational Assessment*, 16(4), 343–364.
- Konvalinka, I., Xygalatas, D., Bulbulia, J. A., Schjødt, U., Jegindø, E.-M., Wallot, S., ... Roepstorff, A. (2011). Synchronized arousal between performers and related spectators in a fire-walking ritual. *Proceedings of the National Academy of Sciences of the United States of America*, 108(20), 8514–9.
- Krahé, C., Springer, A., Weinman, J. A., & Fotopoulou, A. (2013). The social modulation of pain: others as predictive signals of salience - a systematic review. *Frontiers in Human Neuroscience*, 7, 386, 107–127.

- Kreinath, J., Snoek, J., & Stausberg, M. (Eds.). (2006). *Theorizing Rituals, Volume I: Issues, topics, approaches, concepts*. Leiden: Brill.
- Krupp, D. B., DeBruine, L. M., Jones, B. C., & Lalumiere, M. L. (2012). Kin recognition: evidence that humans can perceive both positive and negative relatedness. *Journal of Evolutionary Biology*, 25(8), 1472–1478.
- Kunda, Z. (1987). Motivated inference: Self-serving generation and evaluation of causal theories. *Journal of Personality and Social Psychology*, 53(4), 636–647.
- Lakin, J. L., Jefferis, V. E., Cheng, C. M., & Chartrand, T. L. (2003). The chameleon effect as social glue: Evidence for the evolutionary significance of nonconscious mimicry. *Journal of Nonverbal Behavior*, 27(3), 145–162.
- Lanciano, T., & Curci, A. (2012). Type or dimension? A taxometric investigation of flashbulb memories. *Memory*, 20(2), 177–188.
- Landers, R. (2015). Computing Intraclass Correlations (ICC) as Estimates of Interrater Reliability in SPSS. *The Winnower*, 2:e143518.81744.
- Lang, M., Krátký, J., Shaver, J. H., Jerotijević, D., & Xygalatas, D. (2015). Effects of Anxiety on Spontaneous Ritualized Behavior. *Current Biology*, 25(14), 1892–7.
- Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). International affective picture system (IAPS): Affective ratings of pictures and instruction manual. *Technical Report A-8*.
- Lanman, J. A. (2012). The Importance of Religious Displays for Belief Acquisition and Secularization. *Journal of Contemporary Religion*, 27(1), 49–65.
- Larsen, R. J., & Diener, E. (1992). Promises and problems with the circumplex model of emotion.
- Lawson, E. T., & McCauley, R. N. (1990). *Rethinking Religion: Connecting Cognition and Culture*. Cambridge, UK: Cambridge University Press.
- Leach, C. W., van Zomeren, M., Zebel, S., Vliek, M. L. W., Pennekamp, S. F., Doosje, B., ... Spears, R. (2008). Group-level self-definition and self-investment: a hierarchical (multicomponent) model of in-group identification. *Journal of Personality and Social Psychology*, 95(1), 144–65.
- Legare, C. H., & Herrmann, P. A. (2013). Cognitive consequences and constraints on reasoning about ritual. *Religion, Brain & Behavior*, 3(1), 63–65.
- Legare, C. H., & Souza, A. L. (2012). Evaluating ritual efficacy: Evidence from the supernatural. *Cognition*, 124(1), 1–15.
- Legare, C. H., Watson-Jones, R. E., & Souza, A. L. (2016). A Cognitive Psychological Account of Reasoning about Ritual Efficacy. In H. De Cruz & R. Nichols (Eds.) *Advances in Religion, Cognitive Science, and Experimental Philosophy*. (pp.85-102), London, UK: Bloomsbury.
- Legare, C. H., & Wen, N. J. (2014). The effects of ritual on the development of social group cognition. *International Society for the Study of Behavioural Development*, 66(2), 9–12.
- Legare, C. H., Wen, N. J., Herrmann, P. A., & Whitehouse, H. (2015). Imitative flexibility and the development of cultural learning. *Cognition*, 142, 351–361.
- Leknes, S., & Bastian, B. (2014). The Benefits of Pain. *Review of Philosophy and Psychology*, 5(1), 57–70.
- Leonardelli, G. J., Pickett, C. L., & Brewer, M. B. (2010). Optimal Distinctiveness Theory: A Framework for Social Identity, Social Cognition and Intergroup Relations. *Advances in Experimental Social Psychology*, 43, 63–113.
- Lévi-Strauss, C. (1963). *Totemism*. Boston, NY: Beacon Press.
- Lévi-Strauss, C. (1971). *The Elementary Structures of Kinship*. Boston, NY: Beacon Press.
- Lévi-Strauss, C. (1983). *Mythologiques, Volume I*. Chicago, IL: University of Chicago Press.
- Levine, M., & Crowther, S. (2008). The responsive bystander: how social group membership and group size can encourage as well as inhibit bystander intervention. *Journal of Personality and Social Psychology*, 95(6), 1429–39.
- Lewis, I. M. (1989). *Ecstatic Religion: A Study of Shamanism and Spirit Possession*. Routledge.

- Li, F., & Aksoy, L. (2007). Dimensionality of individualism–collectivism and measurement equivalence of Triandis and Gelfand's scale. *Journal of Business and Psychology*, 21(3), 313–329.
- Lidderdale, J. M., & Walsh, J. J. (1998). The effects of social support on cardiovascular reactivity and perinatal outcome. *Psychology & Health*, 13(6), 1061–1070.
- Lieberman, M. D. (2007). Social cognitive neuroscience: a review of core processes. *Annu. Rev. Psychol.*, 58, 259–289.
- Lieberman, M. D. (2013). *Social: Why our brains are wired to connect*. Oxford, UK: Oxford University Press.
- Liénard, P., & Boyer, P. (2006). Whence Collective Rituals? A Cultural Selection Model of Ritualized Behavior. *American Anthropologist*, 108, 814–827.
- Lin, H.-C., & McFatter, R. (2012). Empathy and distress: two distinct but related emotions in response to infant crying. *Infant Behavior & Development*, 35(4), 887–97.
- Lindemann, B. (2002). The Discovery of Umami. *Chemical Senses*, 27(9), 843–844.
- Lisdorf, A. (2004). The Spread of Non-Natural Concepts. *Journal of Cognition Culture*, 4(1), 151–173.
- Lobetti, T. F. (2013). *Ascetic Practices in Japanese Religion*. London, UK: Routledge.
- LoBue, V., & Rakison, D. H. (2013). What we fear most: A developmental advantage for threat-relevant stimuli. *Developmental Review*, 33(4), 285–303.
- Lock, M. (1993). Cultivating the body: anthropology and epistemologies of bodily practice and knowledge. *Annual Review of Anthropology*, 133–155.
- Lodewijkx, H. F. M. (2005). The Anticipation of a Severe Initiation: Gender Differences in Effects on Affiliation Tendency and Group Attraction. *Small Group Research*, 36(2), 237–262.
- Lodewijkx, H. F. M., & Syroit, J. E. M. M. (1997). Severity of initiation revisited: Does severity of initiation increase attractiveness in real groups? *European Journal of Social Psychology*, 27(3), 275–300.
- López-Martínez, A. E., Esteve-Zarazaga, R., & Ramírez-Maestre, C. (2008). Perceived social support and coping responses are independent variables explaining pain adjustment among chronic pain patients. *The Journal of Pain: Official Journal of the American Pain Society*, 9(4), 373–9.
- Lopez, D. S. (1995). *Curators of the Buddha: The Study of Buddhism under Colonialism*. (D. S. Lopez, Ed.) *Philosophy East and West*. Chicago, IL: University of Chicago Press.
- Lopez, D. S. (1998). *Prisoners of Shangri-La: Tibetan Buddhism and the West*. Chicago, IL: University of Chicago Press.
- Lottrup, L., Stigsdotter, U. K., Meilby, H., & Claudi, A. G. (2015). The workplace window view: a determinant of office workers' work ability and job satisfaction. *Landscape Research*, 40(1), 57–75.
- Lounsbury, M. L., & Bates, J. E. (1982). The Cries of Infants of Differing Levels of Perceived Temperamental Difficultness: Acoustic Properties and Effects on Listeners. *Child Development*, 53(3), 677–686.
- Ludvik, C. (2006). In the Service of the Kaihōgyō Practitioners of Mt. Hiei: The Stopping-Obstacles Confraternity (Sokushō kō) of Kyoto. *Japanese Journal of Religious Studies*, 33, 115–142.
- Luminet, O., & Curci, A. (Eds.). (2008). *Flashbulb Memories: New Issues and New Perspectives*. New York, NY: Psychology Press.
- Lumley, T., Diehr, P., Emerson, S., & Chen, L. (2002). The importance of the normality assumption in large public health data sets. *Annual Review of Public Health*, 23, 151–69.
- Lundqvist, D., Esteves, F., & Öhman, A. (2004). The face of wrath: The role of features and configurations in conveying social threat. *Cognition and Emotion*, 18(2), 161–182.
- Lutz, C., & White, G. M. (1986). The anthropology of emotions. *Annual Review of Anthropology*, 405–436.

- Lyons, D. E., Damrosch, D. H., Lin, J. K., Macris, D. M., & Keil, F. C. (2011). The scope and limits of overimitation in the transmission of artefact culture. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 366(1567), 1158–1167.
- Lyons, D. E., Young, A. G., & Keil, F. C. (2007). The hidden structure of overimitation. *Proceedings of the National Academy of Sciences of the United States of America*, 104(50), 19751–19756.
- Lyubomirsky, S., King, L., & Diener, E. (2005). The benefits of frequent positive affect: does happiness lead to success? *Psychological Bulletin*, 131(6), 803–855.
- MacLean, P. D. (1993). Cerebral evolution of emotion. In M. Lewis, J.M. Haviland-Jones & L.F. Barrett (Eds.) *Handbook of Emotions- Third Edition*. (pp.67–83), New York, NY: Guildford Press
- Mael, F., & Ashforth, B. E. (1992). Alumni and their alma mater: A partial test of the reformulated model of organizational identification. *Journal of Organizational Behavior*, 13, 103–123.
- Malinowski, B. (1966). *Argonauts of the Western Pacific: an account of native enterprise and adventure in the archipelagoes of Melanesia New Guinea*. (7th ed.). London: Routledge & K. Paul.
- Maranise, A. M. J. (2013). Superstition & religious ritual: an examination of their effects and utilization in sport. *Sport Psychologist*, 27(1), 83–91.
- Maratos, F. A., Mogg, K., & Bradley, B. P. (2008). Identification of angry faces in the attentional blink. *Cognition and Emotion*, 22(7), 1340–1352.
- Marshall, D. (2005). Food as ritual, routine or convention. *Consumption Markets & Culture*, 8(1), 69–85.
- Martin, D. (1987). On the Origin and Significance of the Prayer Wheel According to Two Nineteenth-Century Tibetan Literary Sources. *The Journal of the Tibet Society*, 7, i3–29.
- Martin, L. H. (2004). Toward a Scientific History of Religions. In H. Whitehouse & L. H. Martin (Eds.), *Theorizing Religions Past: Archaeology, History, and Cognition*. (pp. 7–14). Walnut Creek, CA: AltaMira Press.
- Master, S. L., Eisenberger, N. I., Taylor, S. E., Naliboff, B. D., Shirinyan, D., & Lieberman, M. D. (2009). A picture's worth: partner photographs reduce experimentally induced pain. *Psychological Science*, 20(11), 1316–8.
- Matthews, M., Lanman, J., Swann, W. B., & Whitehouse, H. (in preparation). Dysphoric experiences in war predict fusion with fellow Veterans.
- McCauley, R. N. (2001). Ritual, memory and emotion: Comparing two cognitive hypotheses. In *Religion in Mind: Cognitive Perspectives on Religious Belief, Ritual and Experience*. (pp. 115–140). Cambridge: Cambridge University Press.
- McCauley, R. N. (2006). How far will an account of ritualized behavior go in explaining cultural rituals? *Behavioral and Brain Sciences*, 29(6), 29–30.
- McCauley, R. N. (2011). *Why religion is natural and science is not*. lecture delivered in May. Oxford, UK: Oxford University Press.
- McCauley, R. N. (2013). Explanatory Pluralism and the Cognitive Science of Religion. In D. Xygalatas & W. W. McCorkle (Eds.), *Mental Culture: Classical Social Theory and the Cognitive Science of Religion*. (pp. 11–32). London, UK: Acumen.
- McCauley, R. N., & Lawson, E. T. (2002). *Bringing Ritual to Mind Psychological Foundations of Cultural Forms*. Cognition. Cambridge: Cambridge University Press.
- McCauley, R. N., & Whitehouse, H. (2005). Introduction: New Frontiers in the Cognitive Science of Religion. *Journal of Cognition and Culture*, 5, 1–13.
- McCutcheon, R. T. (1997). *Manufacturing Religion : The Discourse on Sui Generis Religion and the Politics of Nostalgia*. Oxford, UK: Oxford University Press.

- McDonald, B., & Sylvester, K. (2013). Learning to get drunk: The importance of drinking in Japanese university sports clubs. *International Review for the Sociology of Sport*, 49(3-4), 331–345.
- McFarland, H. N. (1967). *The rush hour of the gods: A study of new religious movements in Japan*. London, UK: Macmillan Pub Co.
- McGuigan, N., Makinson, J., & Whiten, A. (2011). From over-imitation to super-copying: adults imitate causally irrelevant aspects of tool use with higher fidelity than young children. *British Journal of Psychology London England* 1953, 102(1), 1–18.
- McGuire, M. B. (2008). *Lived Religion: Faith and Practice in Everyday Life*. Oxford, UK: Oxford University Press.
- Mead, M. (1943). *Coming of age in Samoa: A study of adolescence and sex in primitive societies*. Harmondsworth, UK: Penguin.
- Mehrabian, A. (1995). Framework for a comprehensive description and measurement of emotional states. *Genetic, Social, and General Psychology Monographs*, 121(3), 339-361.
- Mehrabian, A., & Russell, J. A. (1974). *An approach to environmental psychology*. Cambridge, MA: MIT Press.
- Meier, B. P., Robinson, M. D., & Clore, G. L. (2004). Why Good Guys Wear White: Automatic Inferences About Stimulus Valence Based on Brightness. *Psychological Science*, 15(2), 82–87.
- Merker, B. (2006). Ritual pathology and the nature of ritual culture. *Behavioral and Brain Sciences*, 29(6), 30–31.
- Mesman, J., van IJzendoorn, M. H., & Bakermans-Kranenburg, M. J. (2009). The many faces of the Still-Face Paradigm: A review and meta-analysis. *Developmental Review*, 29(2), 120–162.
- Miller, G. A. (2003). The cognitive revolution: A historical perspective. *Trends in Cognitive Sciences*, 7(3), 141-144.
- Miller, N. (1966). Motives for fear-induced affiliation: emotional comparison or interpersonal similarity?1. *Journal of Personality*, 34(4), 481–503.
- Minowa, Y., Khomenko, O., & Belk, R. W. (2010). Social change and gendered gift-giving rituals: a historical analysis of Valentine's day in Japan. *Journal of Macromarketing*, 31(1), 44-56.
- Miyake, H. (1989). Religious rituals in Shugendo: A summary. *Japanese Journal Of Religious Studies*, 101–116.
- Moeran, B. (1984). *Lost innocence: folk craft potters of Onta, Japan*. Oakland, CA: University of California Press.
- Moeran, B. (2005). Drinking Country: Flows of Exchange in a Japanese Valley. In T.M. Wilson (Ed.) *Drinking Cultures: Alcohol and Identity*. (pp.25–42), New York, NY: Berg.
- Mogg, K., Garner, M., & Bradley, B. P. (2007). Anxiety and orienting of gaze to angry and fearful faces. *Biological Psychology*, 76(3), 163–169.
- Mogg, K., Philippot, P., & Bradley, B. P. (2004). Selective attention to angry faces in clinical social phobia. *Journal of Abnormal Psychology*, 113(1), 160.
- Moghaddam, F. M., & Stringer, P. (1986). Trivial and important criteria for social categorization in the minimal group paradigm. *The Journal of Social Psychology*, 126(3), 345–354.
- Morris, D. B. (1991). *The Culture of Pain*. Oakland, CA: University of California Press.
- Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods*, 7(1), 105.
- Mühlberger, A., Wieser, M. J., & Pauli, P. (2008). Darkness-enhanced startle responses in ecologically valid environments: a virtual tunnel driving experiment. *Biological Psychology*, 77(1), 47–52.
- Mullins, M. R. (2011). Religion in contemporary Japanese lives. In V. Bestor, T. C. Bestor, & A. Yamagata (Eds.), *Routledge Handbook of Japanese Culture and Society*. Routledge, Oxford (pp. 63–74). New York, NY: Routledge.

- Murray, A. D. (1979). Infant crying as an elicitor of parental behavior: an examination of two models. *Psychological Bulletin*, 86(1), 191-215.
- Murray, A. D. (1985). Aversiveness is in the mind of the beholder: Perception of infant caring by adults. In B. M. Lester & C. F. Zachariah Boukydis (Eds.), *Infant crying: Theoretical and research perspectives*. (pp. 217-240). Boston, MA: Springer US.
- Murray, L., & Trevarthen, C. (1986). The infant's role in mother-infant communications. *Journal of Child Language*, 13(1), 15-29.
- Nadel, J., Carchon, I., Kervella, C., Marcelli, D., & Réserbat-Plantey, D. (1999). Expectancies for social contingency in 2-month-olds. *Developmental science*, 2(2), 164-173.
- Nakane, C. (1973). *Japanese Society*. Harmondsworth, UK: Penguin.
- Nelissen, R. M. A. (2011). Guilt-Induced Self-Punishment as a Sign of Remorse. *Social Psychological and Personality Science*, 3(2), 139-144.
- Nelson, J. K. (1996a). *A Year in the Life of a Shinto Shrine*. Seattle, NJ: University of Washington Press.
- Nelson, J. K. (1996b). Freedom of expression: The very modern practice of visiting a Shinto shrine. *Japanese Journal of Religious Studies*, 23(1), 117-153.
- Nelson, J. K. (2000). *Enduring identities: The guise of Shinto in contemporary Japan*. Honolulu, HI: University of Hawaii Press.
- Nelson, J. K. (2008). Household Altars in Contemporary Japan: Rectifying Buddhist“ Ancestor Worship” with Home Décor and Consumer Choice. *Japanese Journal of Religious Studies*, 35(2), 305-330.
- NHK Culture Research Institute. (2008). *2008 NHK survey of religion in Japan- “宗教的なもの” にひかれる日本人～ISSP国際比較調査（宗教）から～*.
- Nielsen, M., Mushin, I., Tomaselli, K., & Whiten, A. (2014). Where Culture Takes Hold: “Overimitation” and Its Flexible Deployment in Western, Aboriginal, and Bushmen Children. *Child Development*, 85(6), 2169-2184.
- Nielsen, M., & Tomaselli, K. (2010). Overimitation in Kalahari Bushman children and the origins of human cultural cognition. *Psychological Science : A Journal of the American Psychological Society / APS*, 21(5), 729-736.
- Norbury, P. (2011). *Japan - Culture Smart!: The Essential Guide to Customs & Culture*. London, UK: Kuperard.
- Norenzayan, A. (2009). Why We Believe: Religion as a Human Universal. In H. Høgh-Olesen (Ed.), *Human Morality and Sociality: Evolutionary and Comparative Perspectives*. (pp. 58-71). New York, NY: Palgrave Macmillan.
- Norenzayan, A. (2013). *Big Gods: How Religion Transformed Cooperation and Conflict*. Princeton, NJ: Princeton University Press.
- Norenzayan, A., Henrich, J., & Slingerland, E. (2013). Religious Prosociality: A Synthesis. In P. J. Richerson & M. H. Christiansen (Eds.), *Cultural Evolution: Science, Technology, Language, and Religion*. (pp. 365-379). Cambridge, MA: MIT Press.
- Norenzayan, A., & Shariff, A. F. (2008). The origin and evolution of religious prosociality. *Science*, 322 (5898), 58-62.
- Nuwer, H. (1990). *Broken Pledges: The Deadly Rite of Hazing*. Atlanta, GA: Longstreet.
- Nuwer, H. (2001). *Wrongs of passage: Fraternities, sororities, hazing, and binge drinking*. Bloomington, IN: Indiana University Press.
- Nuwer, H. (2004). *The Hazing Reader*. Bloomington, IN: Indiana University Press.
- Nwokah, E. E., Davies, P., Islam, A., Hsu, H., & Fogel, A. (1993). Vocal affect in three-year-olds: A quantitative acoustic analysis of child laughter. *The Journal of the Acoustical Society of America*, 94(6), 3076-3090.

- Oakes, P. J., & Turner, J. C. (1980). Social categorization and intergroup behaviour: Does minimal intergroup discrimination make social identity more positive? *European Journal of Social Psychology*, 10(3), 295–301.
- Obeyesekere, G. (1997). *The Apotheosis of Captain Cook European Mythmaking in the Pacific*, Princeton, NJ: Princeton University Press.
- Official Judo Handbook. (2008). Available at: <http://www.usjf.com/downloads-from-usjf/> [Last Accessed 3/3/2016]
- Öhman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: toward an evolved module of fear and fear learning. *Psychological Review*, 108(3), 483–522.
- Öhman, A., & Mineka, S. (2003). The Malicious Serpent snakes as a prototypical stimulus for an evolved module of fear. *Current Directions in Psychological Science*, 12(1), 5–9.
- Olivola, C. Y., & Shafir, E. (2013). The Martyrdom Effect: When Pain and Effort Increase Prosocial Contributions. *Journal of Behavioral Decision Making*, 26(1), 91–105.
- Ortony, A., & Turner, T. J. (1990). What's basic about basic emotions? *Psychological Review*, 97(3), 315.
- Østvik, K., & Rudmin, F. (2001). Bullying and hazing among Norwegian army soldiers: Two studies of prevalence, context, and cognition. *Military Psychology*, 13(1), 17–39.
- Out, D., Pieper, S., Bakermans-Kranenburg, M. J., & van Ijzendoorn, M. H. (2010). Physiological reactivity to infant crying: a behavioral genetic study. *Genes, Brain, and Behavior*, 9(8), 868–76.
- Out, D., Pieper, S., Bakermans-Kranenburg, M. J., Zeskind, P. S., & van Ijzendoorn, M. H. (2010). Intended sensitive and harsh caregiving responses to infant crying: the role of cry pitch and perceived urgency in an adult twin sample. *Child Abuse & Neglect*, 34(11), 863–73.
- Owings, D. H., & Zeifman, D. M. (2004). Human infant crying as an animal communication system: insights from an assessment/management approach. In D.K. Oller & U. Griebel (Eds.) *Evolution of Communication Systems: A Comparative Approach* (pp.151–170), Princeton, MA: MIT Press.
- Oyserman, D., Coon, H. M., & Kemmelmeier, M. (2002). Rethinking individualism and collectivism: evaluation of theoretical assumptions and meta-analyses. *Psychological Bulletin*, 128(1), 3–72.
- Paez, D., Bellelli, G., & Rime, B. (2009). Flashbulb memories, culture and collective memories: Psychosocial processes related to rituals, emotions and memories. In O. Luminet & A. Curci (Eds.), *Flashbulb Memories: New Issues and New Perspectives* (pp. 227–245). New York, NY: Psychology Press.
- Page, R. A., & Moss, M. K. (1976). Environmental influences on aggression: The effects of darkness and proximity of victim. *Journal of Applied Social Psychology*, 6(2), 126–133.
- Pals, D. L. (2006). *Eight theories of religion*. New York: University of Oxford Press.
- Panksepp, J. (1998). *Affective neuroscience: The foundations of human and animal emotions*. Oxford, UK: Oxford university press.
- Panksepp, J. (2007). Neurologizing the psychology of affects: How appraisal-based constructivism and basic emotion theory can coexist. *Perspectives on Psychological Science*, 2(3), 281–296.
- Papa, A., & Bonanno, G. A. (2008). Smiling in the face of adversity: the interpersonal and intrapersonal functions of smiling. *Emotion*, 8(1), 1–12.
- Parkinson, B., Fischer, A. H., & Manstead, A. S. R. (2005). *Emotion in social relations: Cultural, group, and interpersonal processes*. Hove, UK: Psychology Press.
- Parsons, C. E., Young, K. S., Craske, M. G., Stein, A. L., & Kringelbach, M. L. (2014). Introducing the Oxford Vocal (OxVoc) Sounds database: a validated set of non-acted affective sounds from human infants, adults, and domestic animals. *Frontiers in Psychology*, 5, 562.
- Parsons, C. E., Young, K. S., Parsons, E., Stein, A., & Kringelbach, M. L. (2012). Listening to infant distress vocalizations enhances effortful motor performance. *Acta Paediatrica*, 101(4), e189.

- Peligro, K. (2003). *The Gracie Way: An Illustrated History of the World's Greatest Martial Arts Family*. Montpelier, VT: Invisible Cities Press.
- Pershing, J. L. (2006). Men and women's experiences with hazing in a male-dominated elite military institution. *Men and Masculinities*, 8(4), 470–492.
- Peterson, G. R. (2006). Theology and the science wars: Who owns human nature? *Zygon*, 41(4), 853–862.
- Pettas, N. (2012). *Blue Eyed Samurai*. Louisville, KY: Createspace.
- Pinkham, A. E., Griffin, M., Baron, R., Sasson, N. J., & Gur, R. C. (2010). The face in the crowd effect: anger superiority when using real faces and multiple identities. *Emotion*, 10(1), 141.
- Pinter, B., & Greenwald, A. G. (2011). A comparison of minimal group induction procedures. *Group Processes & Intergroup Relations*, 14(1), 81–98.
- Platow, M. J., McClintock, C., & Liebrand, W. (1990). Protecting ontergroup fairness and in-group bias in the minimal group paradigm. *Personality and Social Psychology Bulletin*, 20, 221–239.
- Platow, M. J., Voudouris, N. J., coulson, M., Gilford, N., Jamieson, R., Najdovski, L., Papaleo, N., Pollard, C. Terry, L. (2007). In-group reassurance in a pain setting produces lower levels of physiological arousal: direct support for a self-categorization analysis of social influence. *European Journal of Social Psychology*, 37(4), 649–660.
- Polkinghorne, J. (2001). Some of the Truth. *Science*, 293(5539), 2400.
- Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: an integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology*, 17(3), 715–34.
- Postmes, T., Haslam, S. A., & Jans, L. (2013). A single-item measure of social identification: Reliability, validity, and utility. *British Journal of Social Psychology*, 52(4), 597–617.
- Preacher, K. J., & Kelley, K. (2011). Effect size measures for mediation models: quantitative strategies for communicating indirect effects. *Psychological Methods*, 16(2), 93.
- Preston, S. D., & de Waal, F. B. M. (2002). The communication of emotions and the possibility of empathy in animals. In S.G. Post, L.G. Underwood, J.P. Schloss & W.B. Hurlbut (Eds.) *Altruism and Altruistic Love*. (pp.284–308) Oxford, UK: Oxford University Press.
- Provine, R. R. (1992). Contagious laughter: Laughter is a sufficient stimulus for laughs and smiles. *Bulletin of the Psychonomic Society*, 30(1), 1–4.
- Provine, R. R. (1996a). Contagious Yawning and Laughter: Significance for Sensory Feature Detection. In C.M. Heyes & B.G. Galef (Eds.) *Social Learning in Animals: The Roots of Culture*, (pp.179-203), New York, NY: Academic Press.
- Provine, R. R. (1996b). Laughter. *American Scientist*, 84, 38–45.
- Purzycki, B. G., Apicella, C., Atkinson, Q. D., Cohen, E., McNamara, R. A., Willard, A. K., Xygalatas, D., Norenzayan, A. & Henrich, J. (2016). Moralistic gods, supernatural punishment and the expansion of human sociality. *Nature*, 530(7590), 327-330.
- Purzycki, B. G., & Arakchaa, T. (2013). Ritual Behavior and Trust in the Tyva Republic. *Current Anthropology*, 54(3), 381–388.
- Purzycki, B. G., Haque, O. S., & Sosis, R. (2014). Extending Evolutionary Accounts of Religion beyond the Mind: Religions as Adaptive Systems. In F. Watts & L. P. Turner, (Eds.) *Evolution, Religion and Cognitive Science: Critical and Constructive Essays*. Oxford, UK: Oxford University Press.
- Purzycki, B. G., & Sosis, R. (2013). The extended religious phenotype and the adaptive coupling of ritual and belief. *Israel Journal of Ecology & Evolution*, 59(2), 99–108.
- Purzycki, B. G., & Willard, A. K. (2015). MCI theory: a critical discussion. *Religion, Brain & Behavior*, 1–42.
- Pyysiäinen, I. (2013). Cognitive science of religion: State-of-the-art. *Journal for the Cognitive Science of Religion*, 1(1), 5–28.

- Rabbie, J. M., & Horwitz, M. (1988). Categories versus groups as explanatory concepts in intergroup relations. *European Journal of Social Psychology*, 18(2), 117–123.
- Rabinow, P. (1977). *Reflections on Fieldwork in Morocco*. Berkeley, CA: University of California Press.
- Rakison, D. H., & Derringer, J. (2008). Do infants possess an evolved spider-detection mechanism? *Cognition*, 107(1), 381–393.
- Rakoczy, H., & Tomasello, M. (2009). Done wrong or said wrong? Young children understand the normative directions of fit of different speech acts. *Cognition*, 113(2), 205–212.
- Rakoczy, H., Warneken, F., & Tomasello, M. (2008). The sources of normativity: young children's awareness of the normative structure of games. *Developmental Psychology*, 44(3), 875–881.
- Rakoczy, H., Warneken, F., & Tomasello, M. (2009). Young children's selective learning of rule games from reliable and unreliable models. *Cognitive Development*, 24(1), 61–69.
- Ramey, D. T. (1981). *Group Climate, Campus Image and Attitudes of Fraternity Men Regarding Pledge Hazing*. Bloomington, IN: Indiana University.
- Randolph-Seng, B., & Nielsen, M. E. (2008). Is God really watching you? A response to Shariff and Norenzayan (2007). *The International Journal for the Psychology of Religion*, 18(2), 119–122.
- Ranger, T. O., & Hobsbawm, E. J. (1983). *The invention of tradition*. Cambridge, UK: Cambridge University Press.
- Rappaport, R. A. (1999). *Ritual and Religion in the Making of Humanity*. Cambridge, UK: Cambridge University Press.
- Reader, I. (1990). *Religion in Contemporary Japan*. Basingstoke, UK: MacMillan.
- Reader, I. (1996). *A Poisonous Cocktail?: Aum Shinrikyo's Path to Violence*. Copenhagen: Nordic Institute of Asian Studies Press.
- Reader, I. (2000). *Religious violence in contemporary Japan: The case of Aum Shinrikyō*. Honolulu, HI: University of Hawaii Press.
- Reader, I. (2004a). Dichotomies, Contested Terms and Contemporary Issues in the Study of Religion. *Electronic Journal of Contemporary Japanese Studies*, 1.
- Reader, I. (2004b). Ideology, Academic Inventions and Mystical Anthropology: Responding to Fitzgerald's Errors and Misguided Polemics. *Electronic Journal of Contemporary Japanese Studies*, 3.
- Reader, I. (2013). Reviews of *The Invention of Religion in Japan* by Jason Ānanda Josephson, and: *A Discipline on Foot: Inventing Japanese Native Ethnography 1910–1945* by Alan Christy. *Monumenta Nipponica*, 68(2), 301–313.
- Reader, I., & Tanabe, G. J. (1998). *Practically Religious: Worldly benefits and the common religion of Japan*. Honolulu, HI: University of Hawaii Press.
- Rhodes, R. (1988). John Stevens, The Marathon Monks of Mount Hiei. *Japanese Journal Of Religious Studies*, 16(2-3), 251–3.
- Richerson, P. J., & Boyd, R. (2005). *Not by Genes Alone: How Culture Transformed Human Evolution* (Vol. 52). Chicago, IL: University of Chicago Press.
- Richert, R. A., Whitehouse, H., & Stewart, E. (2005). Memory and analogical thinking in high arousal rituals. In H. Whitehouse & R. N. McCauley (Eds.), *Mind and Religion: Psychological and Cognitive Foundations of Religiosity*. (pp. 127–145). Walnut Creek, CA: AltaMira Press.
- Roemer, M. K. (2010). Shinto festival involvement and sense of self in contemporary Japan. In *Japan Forum*, 22(3-4), 491–512.
- Rosaldo, M. Z. (1980). *Knowledge and passion* (Vol. 4). Cambridge, UK: Cambridge University Press.
- Rosaldo, R. (1980). *Ilongot headhunting, 1883-1974: A study in society and history*. Redwood City, A: Stanford University Press.

- Rosenberg, M. (1987). Depersonalisation: The loss of personal identity. In T. Honess & K. M. Yardley (Eds.), *Self and identity: Perspectives across the lifespan*. (pp. 193–206). London, UK: Routledge & Kegan Paul.
- Rossano, M. J. (2012). The essential role of ritual in the transmission and reinforcement of social norms. *Psychological Bulletin*, 138(3), 529–49.
- Rossen, J. (2013, December 5). The Cult of Lloyd Irvin. *Miami New Times*. Available at: <http://www.miaminewtimes.com/news/the-cult-of-lloyd-irvin-6394125> [Last Accessed 3/3/2016]
- Rugulies, R. (2002). Depression as a predictor for coronary heart disease. *American Journal of Preventive Medicine*, 23(1), 51–61. doi:10.1016/S0749-3797(02)00439-7
- Rupp, K. (2003). *Gift-giving in Japan: Cash, connections, cosmologies*. Redwood City, CA: Stanford University Press.
- Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6), 1161–1178.
- Russell, J. A. (1991). Culture and the categorization of emotions. *Psychological Bulletin*, 110(3), 426–450.
- Russell, J. A. (1994). Is there universal recognition of emotion from facial expressions? A review of the cross-cultural studies. *Psychological Bulletin*, 115(1), 102.
- Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, 110(1), 145–172.
- Russell, J. A. (2009). Emotion, core affect, and psychological construction. *Cognition and Emotion*, 23(7), 1259–1283.
- Russell, J. A., & Barrett, L. F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: dissecting the elephant. *Journal of Personality and Social Psychology*, 76(5), 805–19.
- Russell, J. A., & Carroll, J. M. (1999). On the bipolarity of positive and negative affect. *Psychological Bulletin*, 125(1), 3–30.
- Russell, J. A., & Mehrabian, A. (1977). Evidence for a three-factor theory of emotions. *Journal of Research in Personality*, 11(3), 273–294.
- Russell, J. A., Weiss, A., & Mendelsohn, G. A. (1989). Affect grid: a single-item scale of pleasure and arousal. *Journal of Personality and Social Psychology*, 57(3), 493–502.
- Russell, Y. I., Dunbar, R. I. M., & Gobet, F. (2011). Euphoria versus dysphoria: differential cognitive roles in religion? In S. Masmoudi, A. Naceur, & D. Y. Dai (Eds.), *Attention, Representation & Performance: Integration of Cognition, Emotion & Motivation*. (pp. 147–165). New York, NY: Psychology Press.
- Russell, Y. I., Gobet, F., & Whitehouse, H. (2014). Mood, expertise, analogy, and ritual: an experiment using the five-disk Tower of Hanoi. *Religion, Brain & Behavior*, 1–21.
- Sahlins, M. (1996). *How “Natives” Think: About Captain Cook, for Example*. Chicago, IL: University of Chicago Press.
- Samuel, G. (1993). *Civilized Shamans: Buddhism in Tibetan Societies*. Washington, DC: Smithsonian Institution Press.
- Sander, D., Grafman, J., & Zalla, T. (2003). The human amygdala: an evolved system for relevance detection. *Reviews in the Neurosciences*, 14(4), 303–316.
- Sauter, D. A., Eisner, F., Ekman, P., & Scott, S. K. (2010). Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. *Proceedings of the National Academy of Sciences*, 107(6), 2408–2412.
- Schaller, M., Park, J. H., & Faulkner, J. (2003). Prehistoric dangers and contemporary prejudices. *European Review of Social Psychology*, 14(1), 105–137.

- Schaller, M., Park, J. H., & Mueller, A. (2003). Fear of the dark: Interactive effects of beliefs about danger and ambient darkness on ethnic stereotypes. *Personality and Social Psychology Bulletin*, 29(5), 637–649.
- Schiffmann, R., & Wicklund, R. A. (1992). The Minimal Group Paradigm and its Minimal Psychology: On Equating Social Identity with Arbitrary Group Membership. *Theory Psychology*, 2(1), 29–50.
- Schimmack, U. (2008). The Structure of Subjective Well-Being. In M. Eid & R. J. Larsen (Eds.), *The Science of Subjective Well-Being* (pp. 97–123). New York, NY: Guildford Press.
- Schjødt, U., Sørensen, J., Nielbo, K. L., Xygalatas, D., Mitkidis, P., & Bulbulia, J. A. (2013). Cognitive resource depletion in religious interactions. *Religion, Brain & Behavior*, 3(1), 39–55.
- Schloss, J. P., Barrett, J. L., & Murray, M. J. (2010). Looking Past vs. Overlooking Cognitive–Evolutionary Accounts of Religion: A Response to Nathaniel Barrett. *Journal of the American Academy of Religion*, 78(3), 622–628.
- Schmidt, M. F. H., Rakoczy, H., & Tomasello, M. (2011). Young children attribute normativity to novel actions without pedagogy or normative language. *Developmental Science*, 14(3), 530–539.
- Schneider, A. (2009). The Rhythm of the Whip. *Social Psychology Quarterly*, 72(4), 285–289.
- Seifritz, E., Esposito, F., Neuhoﬀ, J. G., Lüthi, A., Mustovic, H., Dammann, G., Von Bardeleben, U., Radue, E.W., Cirillo, S., Tedeschi, G. & Di Salle, F. (2003). Differential sex-independent amygdala response to infant crying and laughing in parents versus nonparents. *Biological Psychiatry*, 54(12), 1367–1375.
- Sell, J., & Wilson, R. K. (1991). Levels of information and contributions to public goods. *Social Forces*, 70(1), 107–124.
- Semnani-Azad, Z., Sycara, K., & Lewis, M. (2012 May 24th). Dynamics of helping behavior and cooperation across culture. Presented at the *International Conference on Collaboration Technologies and Systems (CTS)*, Denver, CO.
- Shankman, P. (2009). *The trashing of Margaret Mead: anatomy of an anthropological controversy. Studies in American thought and culture*. Madison, WI: University of Wisconsin Press.
- Shariff, A. F., & Norenzayan, A. (2007). God is watching you: priming God concepts increases prosocial behavior in an anonymous economic game. *Psychological Science*, 18(9), 803–809.
- Shariff, A. F., Purzycki, B. G., & Sosis, R. (2014). Religions as cultural solutions to social living. In A.B. Cohen (Ed.) *Culture Reexamined: Broadening Our Understanding of Social and Evolutionary Influences*, (pp. 217–238), Washington, DC: American Psychological Association.
- Shariff, A. F., & Tracy, J. L. (2011). What are emotion expressions for? *Current Directions in Psychological Science*, 20(6), 395–399.
- Shariff, A. F., Willard, A. K., Andersen, T., & Norenzayan, A. (2015). Religious Priming: A Meta-Analysis With a Focus on Prosociality. *Personality and Social Psychology Review*, 20(1), 27–48.
- Shimazono, S. (2004). *From salvation to spirituality: popular religious movements in modern Japan*. Melbourne, Australia: Trans Pacific Press.
- Shiota, M. N., Campos, B., Keltner, D., & Hertenstein, M. J. (2004). Positive emotion and the regulation of interpersonal relationships. *The Regulation of Emotion*, 127–155.
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin*, 86(2), 420–428.
- Shun, I. (1998). The Invention of the Martial Arts: Kano Jigoro and Kodokan Judo. In S. Vlastos (Ed.), *Mirror of Modernity: Invented Traditions of Modern Japan* (pp. 133–143). Berkeley, CA: University of California Press.
- Shweder, R. A. (1993). Everything you ever wanted to know about cognitive appraisal theory without being conscious of it. *Psychological Inquiry*, 4(4), 322–342.
- Silver, M. D., & Brewer, M. B. (1997). Perceptions of Multiple Group Loyalty, and a New Scale for Loyalty Measurement. Presented at the *Annual Convention of the American Psychological Society*. Washington.

- Simon, B., & Hamilton, D. L. (1994). Self-stereotyping and social context: The effects of relative in-group size and in-group status. *Journal of Personality and Social Psychology*, 66, 699–711.
- Singer, T. (2006). The neuronal basis and ontogeny of empathy and mind reading: review of literature and implications for future research. *Neuroscience & Biobehavioral Reviews*, 30(6), 855–863.
- Sleegers, P., Moolenaar, N., Galetzka, M., Pruyn, A., Sarroukh, B., & van der Zande, B. (2012). Lighting affects students' concentration positively: Findings from three Dutch studies. *Lighting Research and Technology*, 45(2), 159–175.
- Slingerland, E., & Bulbulia, J. A. (2011). Evolutionary science and the study of religion. *Religion*, 41(3), 307–328.
- Soh, S., & Leong, F. T. L. (2002). Validity Of vertical and Horizontal Individualism and Collectivism in Singapore Relationships with Values and Interests. *Journal of Cross-Cultural Psychology*, 33(1), 3–15.
- Soler, M. (2012). Costly signaling, ritual and cooperation: evidence from Candomblé, an Afro-Brazilian religion. *Evolution and Human Behavior*, 33(4), 346–356.
- Somnez, C. (2009). slideyfoot.com | bjj resources: Article - BJJ Schools: Size Matters. *slideyfoot.com*, *BJJ resources*. Available at: <http://www.slideyfoot.com/2009/11/bjj-school-size.html> [Last Accessed 3/3/2016]
- Sørensen, J. (2007). Acts that work: A cognitive approach to ritual agency. *Method & Theory in the Study of Religion*, 19(3), 281–300.
- Sørensen, J., Liénard, P., & Feeny, C. (2006). Agent and Instrument in Judgements of Ritual Efficacy. *Journal of Cognition Culture*, 6(3), 463–482.
- Sosis, R. (2005). Does religion promote trust? The role of signaling, reputation, and punishment. *The Interdisciplinary Journal of Research on Religion*, 1(7), 1–30.
- Sosis, R. (2009). The Adaptationist-Byproduct Debate on the Evolution of Religion: Five Misunderstandings of the Adaptationist Program. *Journal of Cognition and Culture*, 9(3), 315–332.
- Sosis, R., & Alcorta, C. S. (2003). Signalling, solidarity, and the sacred: The evolution of religious behavior. *Evolutionary Anthropology*, 12, 264–274.
- Sosis, R., & Bressler, E. R. (2003). Cooperation and Commune Longevity: A Test of the Costly Signaling Theory of Religion. *Cross-Cultural Research*, 37(2), 211–239.
- Sosis, R., Kres, H. C., & Boster, J. S. (2007). Scars for war: evaluating alternative signaling explanations for cross-cultural variance in ritual costs. *Evolution and Human Behavior*, 28(4), 234–247.
- Spelke, E. S. (2000). Core knowledge. *American Psychologist*, 55(11), 1233–1243.
- Spelke, E. S. (2008). The theory of “Core Knowledge.” *Annee Psychologique*, 108(4), 721–756.
- Sperber, D. (1975). *Rethinking Symbolism*. *Cambridge Studies and Papers in Social Anthropology*. Cambridge, UK: Cambridge University Press.
- Sperber, D. (1994). The Modularity of Thought and the Epidemiology of Representations. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the Mind: Domain Specificity in Cognition and Culture* (pp. 39–67). Cambridge, UK: Cambridge University Press.
- Sperber, D. (1996). *Explaining Culture: A Naturalistic Approach*. *Language*. Oxford, UK: Blackwells Publishing.
- Spiro, M. E. (1965). Religious Systems as Culturally constituted Defense Mechanisms. In M. E. Spiro (Ed.), *Concept and Meaning in Cultural Anthropology* (pp. 100–113). New York, NY: The Free Press.
- Spiro, M. E. (1971). Religion: Problems of Definition and Explanation. In M. Banton (Ed.), *Anthropological Approaches to the Study of Religion* (pp. 85–126). London, UK: Tavistock Publications.

- Spiro, M. E. (1982). *Buddhism and society: A great tradition and its Burmese vicissitudes*. Berkeley, NY: University of California Press.
- Spiro, M. E. (1996). Postmodernist Anthropology, Subjectivity, and Science: A Modernist Critique. *Comparative Studies in Society and History*, 38(4), 759–780.
- Spiro, M. E., & D'Andrade, R. G. (1958). A Cross-Cultural Study of Some Supernatural Beliefs. *American Anthropologist*, 60(3), 456–466.
- Spoor, J. R., & Kelly, J. R. (2004). The Evolutionary Significance of Affect in Groups: Communication and Group Bonding. *Group Processes & Intergroup Relations*, 7(4), 398–412.
- Sriram, N., & Greenwald, A. G. (2009). The Brief Implicit Association Test. *Experimental Psychology*, 56(4), 283–294.
- Stallings, J., Fleming, A. S., Corter, C., Worthman, C., & Steiner, M. (2001). The effects of infant cries and odors on sympathy, cortisol, and autonomic responses in new mothers and nonpostpartum women. *Parenting*, 1(1-2), 71–100.
- Stausberg, M. (2014). Fractionizing or Dichotomizing? *Current Anthropology*, 55(6), 688–689.
- Steidle, A., & Werth, L. (2014). In the spotlight: Brightness increases self-awareness and reflective self-regulation. *Journal of Environmental Psychology*, 39, 40–50.
- Stein, D., Apter, A., Ratzoni, G., Har-Even, D., & Avidan, G. (1998). Association between multiple suicide attempts and negative affects in adolescents. *Journal of the American Academy of Child & Adolescent Psychiatry*, 37(5), 488–494.
- Stevens, J. (1988). *The Marathon Monks of Mount Hiei*. Boulder, CO: Shambhala.
- Stormark, K. M., & Braarud, H. C. (2004). Infants' sensitivity to social contingency: A “double video” study of face-to-face communication between 2- and 4-month-olds and their mothers. *Infant Behavior and Development*, 27(2), 195–203.
- Strathern, A. (1970). Male Initiation in New Guinea Highlands Societies. *Ethnology*, 9(4), 373–379.
- Stroeken, K. (2011). Questioning cognitive and interpretive takes on ritual. *Anthropological Theory*, 11(3), 355–372.
- Stryker, S., & Burke, P. J. (2000). The past, present, and future of an identity theory. *Social Psychology Quarterly*, 63(4), 284–297.
- Surakka, V., & Hietanen, J. K. (1998). Facial and emotional reactions to Duchenne and non-Duchenne smiles. *International Journal of Psychophysiology*, 29(1), 23–33.
- Swann, W. B., & Buhrmester, M. D. (2015). Identity Fusion. *Current Directions in Psychological Science*, 24(1), 52–57.
- Swann, W. B., Buhrmester, M. D., Gómez, A., Jetten, J., Bastian, B., Vázquez, A., ... Zhang, A. (2014). What makes a group worth dying for? Identity fusion fosters feelings of familial ties, promoting self-sacrifice. *Journal of Personality and Social Psychology*, 106(6), 912–26.
- Swann, W. B., Gómez, Á., Dovidio, J. F., Hart, S., & Jetten, J. (2010). Dying and killing for one's group: identity fusion moderates responses to intergroup versions of the trolley problem. *Psychological Science*, 21(8), 1176–83.
- Swann, W. B., Gómez, Á., Huici, C., Morales, J. F., & Hixon, J. G. (2010). Identity fusion and self-sacrifice: arousal as a catalyst of pro-group fighting, dying, and helping behavior. *Journal of Personality and Social Psychology*, 99(5), 824–41.
- Swann, W. B., Gómez, Á., Seyle, D. C., Morales, J. F., & Huici, C. (2009). Identity fusion: the interplay of personal and social identities in extreme group behavior. *Journal of Personality and Social Psychology*, 96(5), 995–1011.
- Swann, W. B., Gómez, Á., Vázquez, A., Guillamón, A., Segovia, S., & Carillo, B. (2015). Fusion with the Cross-Gender Group Predicts Genital Sex Reassignment Surgery. *Archives of Sexual Behavior*, 44(5), 1313–8.
- Swann, W. B., Jetten, J., Gómez, A., Whitehouse, H., & Bastian, B. (2012). When group membership gets personal: A theory of identity fusion. *Psychological Review*, 119(3), 441–456.

- Szameitat, D. P., Alter, K., Szameitat, A. J., Wildgruber, D., Sterr, A., & Darwin, C. J. (2009). Acoustic profiles of distinct emotional expressions in laughter. *The Journal of the Acoustical Society of America*, 126(1), 354–366.
- Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American*, 223(5), 96–102.
- Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971a). Social categorisation and intergroup behavior. *Social Psychology of Intergroup Relations*, 1, 149–178.
- Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971b). Social categorization and intergroup behaviour. *European Journal of Social Psychology*, 1(2), 149–178.
- Tajfel, H., & Turner, J. C. (1979). An Integrative Theory of Intergroup Conflict. In W. G. Austin & S. Worchel (Eds.), *The Social Psychology of Intergroup Relations*. (pp. 33–47). Monterey, CA: Brooks/Cole.
- Tajfel, H., & Turner, J. C. (1985). The social identity theory of intergroup behavior. In S. Worchel & W. G. Austin (Eds.) *Psychology of Intergroup Relations 2nd Edition* (pp. 724–747). Chicago, IL: Nelson-Hall.
- Talarico, J. M., & Rubin, D. C. (2003). Confidence, Not Consistency, Characterizes Flashbulb Memories. *Psychological Science*, 14(5), 455–461.
- Talarico, J. M., & Rubin, D. C. (2007). Flashbulb memories are special after all; in phenomenology, not accuracy. *Applied Cognitive Psychology*, 21(5), 557–578.
- Tambiah, S. J. (1976). *World Conqueror and World Renouncer: A study of Buddhism and polity in Thailand against a historical background*. Cambridge, UK: Cambridge University Press.
- Taylor, S. E. (2006). Tend and befriend biobehavioral bases of affiliation under stress. *Current Directions in Psychological Science*, 15(6), 273–277.
- Teichman, Y. (1973). Emotional arousal and affiliation. *Journal of Experimental Social Psychology*, 9(6), 591–605.
- Tellegen, A., Watson, D., & Clark, L. A. (1999). On the Dimensional and Hierarchical Structure of Affect. *Psychological Science*, 10(4), 297–303.
- Thompson, E. R. (2007). Development and Validation of an Internationally Reliable Short-Form of the Positive and Negative Affect Schedule (PANAS). *Journal of Cross-Cultural Psychology*, 38(2), 227–242.
- Tierney, P. (2001). *Darkness in El Dorado: How Scientists and Journalists Devastated the Amazon*. New York, NY: WW Norton & Co..
- Tinbergen, N. (1952). Derived activities; their causation, biological significance, origin, and emancipation during evolution. *The Quarterly Review of Biology*, 27(1), 1–32.
- Tomasello, M. (2008). *Origins of human communication*. Cambridge, MA: MIT Press.
- Tooby, J., & Cosmides, L. (1990). The past explains the present: Emotional adaptations and the structure of ancestral environments. *Ethology and Sociobiology*, 11(4), 375–424.
- Totterdell, P. (2000). Catching moods and hitting runs: mood linkage and subjective performance in professional sport teams. *Journal of Applied Psychology*, 85(6), 848.
- Triandis, H. C., & Gelfand, M. J. (1998). Converging measurement of horizontal and vertical individualism and collectivism. *Journal of Personality and Social Psychology*, 74(1), 118–128.
- Tronick, E. (2003). Things Still To Be Done on the Still-Face Effect. *Infancy*, 4(4), 475–482.
- Tseung-Wong, C. N., & Verkuyten, M. (2010). Intergroup evaluations, group indispensability and prototypicality judgments: A study in Mauritius. *Group Processes & Intergroup Relations*, 13(5), 621–638.
- Turchin, P., Whitehouse, H., Francois, P., Slingerland, E., & Collard, M. (2012). A Historical Database of Sociocultural Evolution. *Cliodynamics: The Journal of Theoretical and Mathematical History*, 3(2).
- Turner, J. C. (1999). Some current issues in research on social identity and self-categorization theories. In N. Ellemers, R. Spears, & B. Doojse (Eds.), *Social Identity* (pp. 6–34). Oxford, UK: Blackwell Publishing.

- Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the Social Group: A Self-Categorization Theory*. Oxford, UK: Blackwell.
- Turner, J. C., & Reynolds, K. J. (2010). The story of social identity. In T. Postmes & N.R. Branscombe (Eds.) *Rediscovering Social Identity: Key Readings* (pp. 13–32). Hove, UK: Psychology Press.
- Turner, R. (2006). Ritual: Meaningful or meaningless? *Behavioral and Brain Sciences*, 29(6), 633.
- Turner, V. (1969). *The Ritual Process: Structure and Anti Structure*. London: Routledge & K. Paul.
- Turner, V. (1974). *Dramas, Fields, and Metaphors: Symbolic Action in Human Society*. Ithaca, N.Y.: Cornell University Press.
- Twigger, R. (1999). *Angry White Pyjamas: An Oxford Poet Trains with the Tokyo Riot Police*. London, UK: Phoenix.
- Upal, M. A. (2010). An alternative account of the minimal counterintuitiveness effect. *Cognitive Systems Research*, 11(2), 194–203.
- van Bunderen, L., & Bastian, B. (2014). “I have paid my dues”: When physical pain reduces interpersonal justice motivations. *Motivation and Emotion*, 38(4), 540–546.
- Van Elk, M., Matzke, D., Gronau, Q. F., Guan, M., Vandekerckhove, J., & Wagenmakers, E.-J. (2015). Meta-analyses are no substitute for registered replications: a skeptical perspective on religious priming. *Frontiers in Psychology*, 6, 1365.
- Van Gennep, A. (1960). *The Rites of Passage. Global Public Health* (Vol. 4). Chicago, IL: Chicago University Press.
- Van Kleef, G. A., Steinel, W., Van Knippenberg, D., Hogg, M. A., & Svensson, A. (2007). Group member prototypicality and intergroup negotiation: how one’s standing in the group affects negotiation behaviour. *British Journal of Social Psychology*, 46(1), 129–152.
- Van Raalte, J. L., Cornelius, A. E., Linder, D. E., Brewer, B. W., Raalte, J. L. Van, Cornelius, A. E., ... Brewer, B. W. (2007). The relationship between hazing and team cohesion. *Journal of Sport Behavior*, 30(4), 491–507.
- Van Slyke, J. A. (2011). *The Cognitive Science of Religion*. London, UK: Ashgate.
- Victoria, B. A. (2006). *Zen at war*. Lanham, MD: Rowman & Littlefield Publishers, Inc.
- Vitebsky, P. (2001). *The Shaman: Voyages of the Soul- Trance, Ecstasy and Healing from Siberia to the Amazon*. London, UK: Duncan Baird.
- Vlastos, S. (1998). *Mirror of Modernity: Invented Traditions of Modern Japan*. Berkeley, CA: University of California Press.
- Waldron, J. J. (2012). A Social Norms Approach to Hazing Prevention Workshops. *Journal of Sport Psychology in Action*, 3(1), 12–20.
- Walker, M. G. (1967). *Organizational Type, Rites of Incorporation, and Group Solidarity: A Study of Fraternity Hell Week*. Seattle, WA: University of Washington Press.
- Walsh, N. E., Schoenfeld, L., Ramamurthy, S., & Hoffman, J. (1989). Normative model for cold pressor test. *American Journal of Physical Medicine & Rehabilitation*, 68(1), 6–11.
- Wang, H., & Bello, P. (2006). Ritualized behavior as a domain-general choice of actions. *Behavioral and Brain Sciences*, 29(6), 633–634.
- Watanabe, N. M. (2015). Sources of Direct Demand: An Examination of Demand for the Ultimate Fighting Championship. *International Journal of Sport Finance*, 10(1), 26–31.
- Watson, D., & Clark, L. A. (1994). The PANAS-X: Manual for the Positive and Negative Affect Schedule - Expanded Form. *Department of Psychology Publications*.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology*, 54(6), 1063–1070.
- Watts, F. (2014). Religion and the emergence of differentiated cognition. In F. Watts & L. P. Turner (Eds.), *Evolution, Religion and Cognitive Science: Critical and Constructive Essays*. Oxford, UK: University of Oxford Press.

- Watts, J. A., Sheehan, O., Atkinson, Q. D., Bulbulia, J., & Gray, R. D. (2016). Ritual human sacrifice promoted and sustained the evolution of stratified societies. *Nature*, 532(7598), 228–231.
- Weber, M. (1930). *The Protestant Work Ethic and the Spirit of Capitalism*. London, UK: George Allen and Unwin.
- Weeden, J., & Kurzban, R. (2013). What predicts religiosity? A multinational analysis of reproductive and cooperative morals. *Evolution and Human Behavior*, 34(6), 440–445.
- Weimann, J. (1994). Individual behaviour in a free riding experiment. *Journal of Public Economics*, 54(2), 185–200.
- Wen, N. J., Herrmann, P. A., & Legare, C. H. (2016). Ritual increases children's affiliation with in-group members. *Evolution and Human Behavior*, 37(1), 54–60.
- Werth, L., Steidle, A., & Hanke, E. (2012). Getting Close in the Dark: Darkness Increases Cooperation. *Proceedings of Experiencing Light*. Available at: Retrieved from <http://2012.experiencinglight.nl/doc/12.pdf> [Last accessed 3/3/2016]
- Whitehouse, H. (1992). Memorable Religions: Transmission, Codification, and Change in Divergent Melanesian Contexts. *Man (New Series)*, 27, 777–797.
- Whitehouse, H. (1995). *Inside the Cult: Religious Innovation and Transmission in Papua New Guinea*. *Oxford studies in social and cultural anthropology*. Oxford, UK: Oxford University Press.
- Whitehouse, H. (1996). Rites of terror: Emotion, Metaphor and Memory in Melanesian initiation cults. *Journal of the Royal Anthropological Institute (New Series)*, 2(4), 703–715. doi:10.2307/3034304
- Whitehouse, H. (2000). *Arguments and Icons: Divergent Modes of Religiosity*. Oxford, UK: Oxford University Press.
- Whitehouse, H. (2002a). Conjectures, Refutations, and Verification: towards a testable theory of “modes of religiosity.” *Journal of Ritual Studies*, 16(2), 44–59.
- Whitehouse, H. (2002b). Modes of Religiosity: Towards a cognitive explanation of the socio-political dynamics of religion. *Method & Theory in the Study of Religion*, 14(3-4), 293–315.
- Whitehouse, H. (2002c). Religious reflexivity and transmissive frequency. *Social Anthropology*, 10(1), 91–103.
- Whitehouse, H. (2004). *Modes of religiosity: A Cognitive Theory of Religious Transmission*. Walnut Creek, CA: AltaMira Press.
- Whitehouse, H. (2008). Cognitive evolution and religion: Cognition and religious evolution. *Issues in Ethnology and Anthropology*, 3(3), 35–47.
- Whitehouse, H. (2013a). Immortality, creation and regulation: updating Durkheim's theory of the sacred. In D. Xygalatas & W. W. McCorkle (Eds.), *Mental Culture: Classical Social Theory and the Cognitive Science of Religion*. (pp. 66–79). London, UK: Acumen.
- Whitehouse, H. (2013b). Ritual and acquiescence to authoritative discourse. *Religion, Brain & Behavior*, 3(1), 76–79.
- Whitehouse, H. (2013c). Three Wishes for the World. *Social Evolution Forum*. Available at <https://evolution-institute.org/focus-article/harvey-whitehouse-three-wishes-for-the-world/> [Last Accessed 3/3/2016]
- Whitehouse, H. (2015). Human rites. *Aeon Magazine*. Available at <https://aeon.co/essays/rituals-define-us-in-fathoming-them-we-might-shape-ourselves> [Last Accessed 3/3/2016]
- Whitehouse, H., & Hodder, I. (2010). Modes of religiosity at Çatalhöyük. In I. Hodder (Ed.), *Religion in the Emergence of Civilization: Çatalhöyük as a Case Study* (pp. 122–145). Cambridge, UK: Cambridge University Press.
- Whitehouse, H., Jong, J., Buhrmester, M. D., Gómez, A., Bastian, B., Kavanagh, C. M., Newson, Matthews, M., Lanman, J., & Gavrillets, S. (submitted). Closer than brothers: The evolutionary roots of identity fusion and self-sacrifice. *PLoS Biology*.

- Whitehouse, H., & Laidlaw, J. (2004). *Ritual and Memory: Towards a Comparative Anthropology of Religion*. (H. Whitehouse & R. N. McCauley, Eds.). Walnut Creek, CA: AltaMira Press.
- Whitehouse, H., & Laidlaw, J. (2007). *Religion, Anthropology and Cognitive Science*. Durham: Carolina Academic Press.
- Whitehouse, H., & Lanman, J. A. (2014). The ties that bind us: Ritual, fusion and identification. *Current Anthropology*, 55, 674–695.
- Whitehouse, H., & Legare, C. H. (2012). The imitative foundations of social learning. In *Annual Meeting of the Human Behavior and Evolution Conference*. Aubuerque, NM.
- Whitehouse, H., & Martin, L. H. (2004). *Theorizing religions past: Archaeology, history and cognition*. Walnut Creek, CA: AltaMira Press.
- Whitehouse, H., & McCauley, R. N. (2005). *Mind and Religion: The Psychological and Cognitive Foundations of Religiosity*. Walnut Creek, CA: AltaMira Press.
- Whitehouse, H., & McQuinn, B. (2012). Divergent Modes of Religiosity and Armed Struggle. In M. Juergensmeyer, M. Kitts, & M. Jerryson (Eds.), *The Oxford Handbook of Religion and Violence* (pp. 597–619). Oxford, UK: Oxford University Press.
- Whitehouse, H., McQuinn, B., Buhrmester, M. D., & Swann, W. B. (2014). Brothers in arms: Libyan revolutionaries bond like family. *Proceedings of the National Academy of Sciences of the United States of America*, 111(50), 17783–5.
- Whiten, A., Horner, V., & Marshall-Pescini, S. (2005). Selective Imitation in Child and Chimpanzee: A Window on the Construal of Others' Actions. In S. Hurley & N. Chater (Eds.), *Perspectives on imitation From cognitive neuroscience to social science Volume 1* (Vol. 1, pp. 263–283). Cambridge, MA: MIT Press.
- Whiten, A., McGuigan, N., Marshall-Pescini, S., & Hopper, L. M. (2009). Emulation, imitation, over-imitation and the scope of culture for child and chimpanzee. *Philosophical Transactions of the Royal Society of London - Series B: Biological Sciences*, 364(1528), 2417–2428.
- Wiebe, D. (2004). Critical Reflections on the Modes of Religiosity Argument. In H. Whitehouse & L. H. Martin (Eds.), *Theorizing Religions Past: Archaeology, History, and Cognition*. (pp. 197–213). Walnut Creek, CA: AltaMira Press.
- Wiech, K., Farias, M., Kahane, G., Shackel, N., Tiede, W., & Tracey, I. (2008). An fMRI study measuring analgesia enhanced by religion as a belief system. *Pain*, 139(2), 467–76.
- Wierzbicka, A. (1986). Human emotions: universal or culture-specific? *American Anthropologist*, 88(3), 584–594.
- Wildman, W. J., & Sosis, R. (2011). Stability of Groups with Costly Beliefs and Practices. *Journal of Artificial Societies and Social Simulation*, 14(3), 6.
- Wilson, D. S. (2010). *Darwin's Cathedral: Evolution, Religion, and the Nature of Society*. Chicago, IL: University of Chicago Press.
- Wiltermuth, S. S., & Heath, C. (2009). Synchrony and cooperation. *Psychological Science*, 20(1), 1–5.
- WIN-Gallup International. (2012). *Global Index of Religiosity and Atheism*.
- WinGallup. (2014). *End of Year Survey 2014- Japan*.
- Winograd, E., & Neisser, U. (2006). *Affect and Accuracy in Recall: Studies of "Flashbulb" Memories* (Vol. 2). Cambridge, UK: Cambridge University Press.
- Winslow, D. (1999). Rites of Passage and Group Bonding in he Canadian Airborne. *Armed Forces & Society*, 25(3), 429–457.
- Wirth, M. M., & Schultheiss, O. C. (2006). Effects of affiliation arousal (hope of closeness) and affiliation stress (fear of rejection) on progesterone and cortisol. *Hormones and Behavior*, 50(5), 786–95.
- Xygalatas, D. (2012). *The Burning Saints: Cognition and Culture in the Fire-walking Rituals of the Anastenaria*. Sheffield: Equinox Publishing Ltd.

- Xygalatas, D. (2013a). Effects of religious setting on cooperative behavior: A case study from Mauritius. *Religion, Brain & Behavior*, 3(2), 91–102.
- Xygalatas, D. (2013b). What is natural and unnatural about religion and science? *Religion, Brain & Behavior*, 3(2), 161–164.
- Xygalatas, D. (2014). Ritual and cohesion. What is the place of euphoric arousal?, *Current Anthropology*, 55(6).
- Xygalatas, D., Konvalinka, I., Bulbulia, J. A., & Roepstorff, A. (2011). Quantifying collective effervescence: Heart-rate dynamics at a fire-walking ritual. *Communicative & Integrative Biology*, 4(6), 735–738.
- Xygalatas, D., Schjødt, U., Bulbulia, J. A., Konvalinka, I., Jegindø, E.-M., Reddish, P., Geertz, A.W. & Roepstorff, A. (2013a). Autobiographical Memory in a Fire-Walking Ritual. *Journal of Cognition and Culture*, 13(1-2), 1–16.
- Xygalatas, D., Mitkidis, P., Fischer, R., Reddish, P., Skewes, J., Geertz, A. W., Roepstorff, A. & Bulbulia, J. A. (2013b). Extreme rituals promote prosociality. *Psychological Science*, 24(8), 1602–5.
- Yamagishi, T., Mifune, N., Liu, J. H., & Pauling, J. (2008). Exchanges of group-based favours: Ingroup bias in the prisoner's dilemma game with minimal groups in Japan and New Zealand. *Asian Journal of Social Psychology*, 11(3), 196–207.
- Yoshida, F., & Kubota, K. (1994). Minority versus majority: intergroup discrimination in the minimal group paradigm. *Shinrigaku Kenkyu The Japanese Journal of Psychology*, 65(5), 346–354.
- Yoshihara, C., & Shimizu, S. (2005). Social networks in drinking behaviors among Japanese: support network, drinking network, and intervening network. *Nihon Arukoru Yakubutsu Igakkai Zasshi Japanese Journal of Alcohol Studies Drug Dependence*, 40(5), 441–454.
- Young, F. W. (1965). *Initiation ceremonies: a cross-cultural study of status dramatization*. Indianapolis, NY: The Bobbs-Merrill Co., Inc.
- Younger, J., Aron, A., Parke, S., Chatterjee, N., & Mackey, S. (2010). Viewing pictures of a romantic partner reduces experimental pain: involvement of neural reward systems. *PloS One*, 5(10), e13309.
- Yuki, M. (2003). Intergroup Comparison Versus Intragroup Relationships : A Cross-Cultural Examination of Social Identity Theory in North American and East Asian Cultural Contexts. *Social Psychology Quarterly*, 66(2), 166–183.
- Yuki, M. (2011). Intragroup Relationships and Intergroup Comparisons as two sources of Group-Based Collectivism. In *Social Cognition, Social Identity, and Intergroup Relations: A Festschrift in honor of Marilynn B. Brewer*. (pp. 247–266). Philadelphia PA: Psychology Press.
- Yuki, M., & Schug, J. (2012). Relational mobility: A socioecological approach to personal relationships. In O. Gillath, G. Adams, & A. Kunkel (Eds.), *Relationship Science: Integrating Evolutionary, Neuroscience, and Sociocultural Approaches*. (pp. 137–151). Washington, DC: American Psychological Association.
- Yuki, M., & Takemura, K. (2014). Intergroup comparison and intragroup relationships: Group processes in the cultures of individualism and collectivism. In M. Yuki & M. Brewer (Eds.), *Culture and Group Processes* (pp. 38–65). Oxford, UK: Oxford University Press.
- Yuki, M., & Yokota, K. (2009). The primal warrior: Outgroup threat priming enhances intergroup discrimination in men but not women. *Journal of Experimental Social Psychology*, 45(1), 271–274.
- Zadeh, R. S., Shepley, M. M., Williams, G., & Chung, S. S. E. (2014). The impact of windows and daylight on acute-care nurses' physiological, psychological, and behavioral health. *HERD: Health Environments Research & Design Journal*, 7(4), 35–61.
- Zahavi, A. (1975). Mate selection-a selection for a handicap. *Journal of Theoretical Biology*, 53(1), 205–14.

- Zahavi, A., & Zahavi, A. (1997). *The Handicap Principle: A Missing Piece of Darwin's Puzzle*. Oxford, UK: Oxford University Press.
- Zajonc, R. B. (1985). Emotion and facial efference: A theory reclaimed. *Science*, 228(4695), 15–21.
- Zegans, S., & Zegans, L. S. (1978). Bar Mitzvah: A rite for a transitional age. *Psychoanalytic Review*, 66(1), 115–132.
- Zeifman, D. M. (2001). An ethological analysis of human infant crying: Answering Tinbergen's four questions. *Developmental Psychobiology*, 39(4), 265–285.
- Zeskind, P. S. (1983). Cross-Cultural Differences in Maternal Perceptions of Cries of Low- and High-Risk Infants. *Child Development*, 54(5), 1119–1128.
- Zeskind, P. S. (1985). A developmental perspective of infant crying. In *Infant crying* (pp. 159–185). Springer.
- Zeskind, P. S. (1987). Adult heart-rate responses to infant cry sounds. *British Journal of Developmental Psychology*, 5(1), 73–79.
- Zeskind, P. S., & Barr, R. G. (1997). Acoustic characteristics of naturally occurring cries of infants with "colic." *Child Development*, 68(3), 394–403.
- Zeskind, P. S., Klein, L., & Marshall, T. R. (1992). Adults' perceptions of experimental modifications of durations of pauses and expiratory sounds in infant crying. *Developmental Psychology*, 28(6), 1153.
- Zeskind, P. S., & Lester, B. M. (1978). Acoustic features and auditory perceptions of the cries of newborns with prenatal and perinatal complications. *Child Development*, 580–589.
- Zeskind, P. S., & Marshall, T. R. (1988). The relation between variations in pitch and maternal perceptions of infant crying. *Child Development*, 193–196.
- Zimmerman, M. (1985). *How to Do Business with the Japanese*. Random House.