Social bonding through processing fluency for interpersonal synchronisation



Joshua Michael Silberstein Bamford

St Peter's College

Institute of Human Sciences

University of Oxford

A thesis submitted for the degree of Doctor of Philosophy.

Supervised by Prof. Emma Cohen and Dr Bronwyn Tarr.

Trinity, 2022

Final version: 12/12/2022

Acknowledgements

It has been an immense privilege to work on a project that brings together my interests in music, cognitive science, and evolution. First and foremost, I wish to thank Emma Cohen and Bronwyn Tarr for agreeing to supervise me in this project, and for continuing to inspire me with your rigour as academics and kindness as people.

Thank you to my parents, Mike and Mandy Bamford, for being my biggest supporters throughout my academic journey and for being my first science teachers. I also wish to thank my wife, Fabienne, for accompanying me on much of that journey and for her proof-reading skills which are second to none. To all my family and friends around the world, I'm sorry for my absence, but I'm grateful you have stuck around anyway, and I hope to see you soon.

I've been lucky to have some tremendously supportive colleagues. Lunches at the Institute of Human Sciences in Oxford, and at the School of Psychological Science in Perth, were the only thing keeping me going at times. Thank you to all of my fellow students and researchers at those institutes for providing your support, statistics help, life advice, and participation in pilot studies. Particular thanks must go to Lynden Miles at UWA, for taking me into his lab, and to Cathy Macpherson and Bradley Walker for their assistance. I'm especially grateful to everyone in the Social Body Lab in Oxford, particularly Emma and Bronwyn, along with Sophie Berdugo, Arran Davis, Jon Gorrie and Bahar Tunçgenç. One might expect that a research group that studies social bonding would be a friendly place to work, but that wouldn't do justice to the supportive environment they cultivated here.

I am grateful to all the academics with whom I have shared stimulating, inspiring, frustrating, and illuminating conversations about my research, many of whom are cited within this thesis. This includes, but is certainly not limited to: Robin Dunbar, Eiluned Pearce, Laura Bishop, Jan Stupacher, Emily Carlson, Birgitta Burger, Petri Toiviainen, Jonna Vuoskoski, Daniel Bowling, Lisa Horn, Tecumseh Fitch, Jane Davidson, Peter Keller, Susana Carvalho, Martha Newson, Maxwell Ramstead, Samuel Veissière, Eric Clarke, and Patrick Savage. Science can be like standing on the shoulders of giants, but it can also feel like a network of friends. I'd particularly like to mention Iain Morley and Jeremy Montagu, and to express great sorrow that our conversations were cut tragically short. My deepest gratitude also to Nicholas Bannan, ever my friend and mentor, since he first encouraged me down this path of research.

I wish to thank all those who distracted me from my thesis as well: the St Peter's College MCR, St Peter's College Choir, Wadham College Choir, Magdalen Players, Laurelin Voices, Oxford Opera Society, Oxford University Tea Appreciation Society, Oxford Lindy Hoppers, Oxford JSoc (although honourable mention to them for feeding me every Friday), the SysMus Executive Committee, and OxBAT organising committee. You impeded my progress but kept me sane.

Finally, thank you to my grandparents, to whom I wish to dedicate this thesis. To Wendy Silberstein, who was simultaneously the most determined, curious and kindest person I've ever met, and to Peter Silberstein, for first piquing my interest in the evolution of the nervous system and for lending me his shoes.

Abstract

It has often been observed that people like each other more after synchronising actions. This is referred to as the synchrony-bonding effect and it has been observed in a wide variety of settings – from dance to rowing. It has been suggested that music may have evolved alongside human rhythmic abilities in order to support synchronised action to promote social bonding and group cohesion. However, the precise cognitive mechanisms underpinning the synchrony-bonding effect remain opaque.

This thesis aims to bring new insights to the synchrony-bonding effect by testing a novel hypothesis: that the social bonding effects of synchrony arise from processing fluency. To do so, I bring together the general literature on processing fluency as it applies to perceptual preferences and aesthetic judgements, together with existing studies of interpersonal synchrony. I then conduct three empirical studies using novel experimental designs to test the fluency through synchrony hypothesis. Study 1 found that synchrony reduces processing load on a simple visual perception task, when compared with non-synchrony. Study 2 found that synchrony reduces processing load when participating in a rhythmic tapping task, when compared with non-synchrony. Study 3 built upon Study 2 and found that perceived difficulty of a rhythmic tapping task mediated the social bonding effects of synchrony.

These studies reframe synchrony as a means for reducing processing load, which could explain the synchrony-bonding effect as being an effect of processing fluency, rather than of synchrony *per se*. Thus, the act of synchronising may be considered an act of making oneself easier to process. I discuss these findings in the context of the broader synchrony-bonding literature, consider the implications for the evolution of rhythmic abilities, and attempt to highlight the importance of understanding synchrony in an increasingly de-synchronised world.

Glossary

AIF – Active Inference Framework

ASA – Auditory Scene Analysis

ANOVA – Analysis of Variance

BFI-10 – Big Five Index 10-item version

BPM – Beats per minute

DC – Direction Congruent

DI – Direction Incongruent

FEP – Free Energy Principle

Gold-MSI – Goldsmiths Musical Sophistication Index

IOI – Inter-Onset Interval

IPA – Index of Pupillary Activity

LHIPA – Low/High Index of Pupillary Activity

LMEM – Linear Mixed Effects Model

PIA – Prioritised Integrative Attending

RM-ANOVA – Repeated-Measures Analysis of Variance

RT – Reaction Time

SC – Speed Congruent

SI – Speed Incongruent

Table of Contents

ACKNOWLEDGEMENTS ..... **II**

ABSTRACT ..... **III**

GLOSSARY ..... **IV**

TABLE OF CONTENTS ..... **V**

LIST OF FIGURES ..... **VIII**

LIST OF TABLES ..... **IX**

STATEMENT OF CONTRIBUTIONS ..... **X**

0 INTRODUCTION **11**

0.1 THE SYNCHRONY-BONDING EFFECT 12
0.1.1 Mechanisms 13
0.2 FLUENCY THROUGH SYNCHRONY 16
0.2.1 Chapter 1 - Theory 18
0.2.2 Chapter 2 – 'Ease of perception in synchrony' (Study 1) 19
0.2.3 Chapter 3 – 'Ease of participation in synchrony' (Study 2) 20
0.2.4 Chapter 4 – 'Synchrony-bonding through fluency' (Study 3) 21
0.2.5 Chapter 5 – General discussion 22
0.3 SUMMARY 23

1 THE ROLE OF PROCESSING FLUENCY IN THE SYNCHRONY-BONDING EFFECT **25**

1.1 INTRODUCTION 25
1.2 SYNCHRONY 26
1.2.1 The synchrony-bonding effect 27
1.3 PROCESSING FLUENCY 29
1.3.1 Fluency factors 31
1.3.2 Context 33
1.3.3 Stimulus features 34
1.3.4 Prior experience 36
1.3.5 Fluency of Production 40
1.3.6 Competing drives for fluency and novelty 41
1.4 FLUENCY IN SYNCHRONY 44
1.4.1 Ease of attention 47
1.4.2 Ease of anticipation 51
1.4.3 Ease of adaptation 54
1.4.4 Integrating attention, anticipation and adaptation 55
1.5 SUMMARY 56
1.6 ACKNOWLEDGEMENTS 60

2 PROCESSING FLUENCY FOR VISUAL SYNCHRONY PERCEPTION **61**

2.1 INTRODUCTION 61
2.1.1 Processing Fluency 62
2.1.2 The Present Study 65
2.2 STUDY 1A 68
2.2.1 Participants 68
2.2.2 Materials 69
2.2.3 Procedure 70
2.2.4 Analysis 71
2.2.5 Results 72

2.2.6 Discussion 75
2.3 Study 1b 78
2.3.1 Participants 78
2.3.2 Materials 79
2.3.3 Procedure 80
2.3.4 Analysis 80
2.3.5 Results 81
2.3.6 Discussion 85
2.4 General Discussion 87
2.5 Acknowledgements 94

3 SYNCHRONY REDUCES COGNITIVE LOAD DURING A CONTINUOUS TAPPING TASK 95

3.1 Introduction 95
3.1.1 Dual tasks and auditory distraction 97
3.1.2 Pupil measures of processing load 98
3.1.3 Fluency through synchrony 99
3.2 Method 101
3.3 Results 108
3.4 Discussion 114
3.5 Conclusion 120
3.6 Acknowledgements 120

4 SYNCHRONY-BONDING EFFECT MEDIATED BY PERCEIVED DIFFICULTY 122

4.1 Introduction 122
4.1.1 Temporal and timbral coherence 124
4.1.2 Measuring fluency and bonding 129
4.1.3 Fluency-through-synchrony 132
4.2 Method 133
4.3 Results 139
4.3.1 Effect of tempo ratio on bonding and processing load 140
4.3.2 Effect of timbral similarity on bonding 142
4.3.3 Mediation of synchrony-bonding effect by processing load 144
4.4 Discussion 145
4.4.1 Effect of Tempo Ratio on Bonding (H1) 145
4.4.2 Effect of Tempo Ratio on cognitive load (H2) 146
4.4.3 Timbral similarity (H3 and H4) 148
4.4.4 Bonding mediated by processing load (H5 and H6) 148
4.4.5 The 120-80 BPM condition 149
4.5 General Discussion 150
4.6 Acknowledgements 154

5 GENERAL DISCUSSION 155

5.1 Synchrony-bonding theories 156
5.2 Fluency through synchrony 159
5.3 Integration between theories 161
5.4 Deeper mechanisms 165
5.4.1 Attention and streaming 165
5.4.2 Anticipation and prediction 166
5.5 Constraints and Limitations 169
5.6 Further implications and future research 172
5.6.1 Individual differences 173
5.6.2 Rhythm ratios and optimum complexity 173
5.6.3 Multiscale synchrony 174

5.6.4 Multilevel selection 177
5.6.5 The future of synchrony 179
5.7 FINAL CONCLUSIONS 180

REFERENCES 183

6 APPENDICES 207

6.1 APPENDICES FOR CHAPTER 2 207
6.1.1 Participant information and consent form 207
6.1.2 Demographic questions 211
6.2 APPENDICES FOR CHAPTER 3 212
6.2.1 Advertisement 212
6.2.2 Participant information sheet 213
6.2.3 Participant consent form 215
6.2.4 Instructions for participants 216
6.2.5 Debrief sheet 216
6.2.6 Big Five Index – 10 Item version (BFI-10) 218
6.2.7 Goldsmiths Musical Sophistication Inventory (Gold-MSI) 218
6.2.8 Demographic questions 220
6.2.9 Supplementary Analysis 221
6.2.10 Video stimuli 222
6.3 APPENDICES FOR CHAPTER 4 222
6.3.1 Advertisement 222
6.3.2 Participant information sheet 223
6.3.3 Participant consent form 225
6.3.4 Instructions for participants 226
6.3.5 Post-task questions 227
6.3.6 Debrief sheet 227
6.3.7 Big Five Index – 10 Item version (BFI-10) 229
6.3.8 Goldsmiths Musical Sophistication Inventory (Gold-MSI) 229
6.3.9 Demographic questions 231
6.3.10 Supplementary analyses 232
6.3.11 Video stimuli 234

List of Figures

FIGURE 1.1 INFORMATION PROCESSING MODEL LEADING TO AESTHETIC APPRAISAL FOR A WORK OF ART (LEDER, 2013). 32
FIGURE 1.2 FACTORS THAT AFFECT INTERPERSONAL COORDINATION (KELLER ET AL., 2014). 46
FIGURE 1.3 VISUAL BISTABILITY, IN WHICH THE IMAGE MAY BE SEEN EITHER AS TWO FACES OR ONE VASE (PRESSNITZER ET AL., 2011). 49
FIGURE 2.1 A SINGLE GABOR PATCH. 66
FIGURE 2.2 ARRANGEMENT OF STIMULI IN THE GABOR-FLANKER TASK. THE CENTRAL GABOR WAS ALWAYS THE TARGET STIMULUS, AND THE PATCHES IMMEDIATELY ABOVE AND BELOW IT WERE ALWAYS STATIC, WHILE THOSE TO ITS LEFT AND RIGHT WERE ANIMATED. THE FLANKER STIMULI EITHER MOVED IN THE SAME DIRECTION AS THE TARGET (CONGRUENT) OR IN THE OPPOSITE DIRECTION (INCONGRUENT). ARROWS FOR ILLUSTRATION PURPOSES ONLY. 70
FIGURE 2.3 REACTION TIME (S) ON CORRECT TRIALS, AND ACCURACY (%) PER BLOCK, OVER FLANKER CONDITION (STIMULUS TYPE). INCLUDES OUTLIERS, SHOWING THOSE WITH LOW ACCURACY ON INCONGRUENT TRIALS. 73
FIGURE 2.4 MEAN REACTION TIME (MS) ON CORRECT TRIALS PER BLOCK, PLOTTED AGAINST ACCURACY (%) PER BLOCK. 73
FIGURE 2.5 PLOT OF EACH PARTICIPANT'S COMPOSITE PERFORMANCE SCORE (ACCURACY/REACTION TIME) IN BOTH THE CONGRUENT AND INCONGRUENT CONDITIONS, AGAINST PRESENTATION ORDER (BLOCK NUMBER). ALL DATA ARE PRESENTED, INCLUDING SUBJECTS THAT SCORED BELOW CHANCE LEVEL ACCURACY. ERROR BARS SHOW 95% CONFIDENCE INTERVALS. 75
FIGURE 2.6 MEAN REACTION TIME, ACCURACY AND PERFORMANCE PER CONDITION. EACH POINT REPRESENTS A PARTICIPANT PER CONDITION, AVERAGED ACROSS TRIALS PER CONDITION. 82
FIGURE 2.7 INTERACTION BETWEEN THE SPEED AND DIRECTION FLANKER CONDITIONS. ERROR BARS SHOW 95% CONFIDENCE INTERVALS. 84
FIGURE 2.8 INTERACTION BETWEEN TARGET SPEED AND SPEED CONGRUENCY. ERROR BARS SHOW 95% CONFIDENCE INTERVALS. 85
FIGURE 2.9 THE CONDITIONS FROM STUDY 1B (BLUE REPRESENTING THE TARGET AND ORANGE REPRESENTING THE FLANKERS), PLOTTED AS LINEAR FUNCTIONS SHOWING CHANGE IN POSITION OVER TIME. 89
FIGURE 2.10 ILLUSTRATION OF TWO OSCILLATORS (ONE IN BLUE, THE OTHER IN ORANGE) MOVING OVER TIME IN: (A) SYNCHRONY, (B) ANTI-PHASE SYNCHRONY, (C) NON-SYNCHRONY. 90
FIGURE 3.1 SCREEN CAPTURE DURING THE EXPERIMENT. THE PARTICIPANT WAS ASKED TO TAP IN TIME WITH THE TARGET ON THE LEFT. IN THIS INSTANCE, THE DISTRACTOR WAS IN THE STILL CONDITION. HERE A SINGLE TRIAL OF THE LETTER SEARCH TASK APPEARS SUPERIMPOSED ON TOP OF THE VIDEO STIMULUS. 102
FIGURE 3.2 EXPERIMENT DESIGN SHOWING THE SIX BLOCKS, EACH OF 50 TRIALS IN THE LETTER SEARCH TASK, AND WITH AN EXAMPLE OF THE THREE CONDITIONS FOR THE TAPPING TASK IN RANDOMISED ORDER, PRESENTED TWICE. 105
FIGURE 3.3 VIOLIN PLOTS FOR PERFORMANCE (ER/RT), LHIPA, AVERAGED PUPIL SIZE, NUMBER OF FIXATIONS, IOI VARIANCE, AND MEAN IOI BY CONDITION. INTERNAL BOXPLOT SHOWS MEDIAN, FIRST QUARTILE, AND 1.5 QUARTILE. 111
FIGURE 3.4 LINEAR MIXED-EFFECTS MODEL FOR PERFORMANCE BY CONDITION AND IOI VARIANCE. THE SHADED AREA REPRESENTS 95% CONFIDENCE INTERVALS AROUND THE REGRESSION LINE. 113
FIGURE 4.1 SCREENSHOT OF A VIDEO STIMULUS. PARTICIPANTS WERE INSTRUCTED TO DRUM ALONG IN TIME WITH THE PERSON ON THE LEFT OF THE SCREEN. 134
FIGURE 4.2 VIOLIN PLOTS OF EFFECT OF TEMPO RATIO ON PERCEIVED DIFFICULTY, BONDING, PUPIL SIZE, LHIPA, IOI VARIANCE, AND MEAN IOI. 141
FIGURE 4.3 LINEAR MIXED-EFFECTS MODEL PLOTTING THE EFFECTS OF TEMPO RATIO AND TIMBRE CONDITIONS (SAME OR DIFFERENT) ON SELF-REPORTED SOCIAL BONDING. 143
FIGURE 4.4 MEDIATION MODEL SHOWING THE DIRECT, INDIRECT AND TOTAL EFFECT OF TEMPO RATIO ON BONDING, AS MEDIATED BY DIFFICULTY. TEMPO RATIO IS SPLIT INTO TWO FACTORS COMPARING COMPLEMENTARY AND NON-SYNCHRONY TO SYNCHRONY (THE REFERENCE CONDITION). (* $p < .05$, ** $p < .01$, *** $p < .001$) 144
FIGURE 4.5 MEDIATION MODEL SHOWING THE DIRECT, INDIRECT AND TOTAL EFFECT OF TEMPO RATIO ON BONDING, AS MEDIATED BY PUPIL SIZE. TEMPO RATIO IS SPLIT INTO TWO FACTORS COMPARING COMPLEMENTARY AND NON-SYNCHRONY TO SYNCHRONY (THE REFERENCE CONDITION) 145
FIGURE 5.1 PROPOSED MODEL OUTLINING THE POSSIBLE RELATIONSHIP BETWEEN DIFFERENT SYNCHRONY-BONDING THEORIES 164
FIGURE 6.1 INTERACTION EFFECT OF PUPIL SIZE AND DIFFICULTY ON BONDING. DIFFICULTY WAS DIVIDED INTO 5 BANDS, EACH CORRESPONDING TO THE COLOURS IN THE LEGEND. SHADING SHOWS THE CONFIDENCE INTERVAL OF THE MODEL. 234

List of Tables

TABLE 2.1 DESCRIPTIVE STATISTICS FOR REACTION TIME, ACCURACY, AND COMPOSITE PERFORMANCE ACROSS THE TWO STIMULI CONDITIONS, WITH OUTLIERS EXCLUDED. 72
TABLE 2.2 REPEATED-MEASURES ANOVA FOR THE EFFECT OF STIMULUS CONDITION AND BLOCK NUMBER ON PERFORMANCE. 74
TABLE 2.3 DESCRIPTIVE STATISTICS FOR MEAN REACTION TIME, ACCURACY AND PERFORMANCE ACROSS ALL CONDITIONS ON BOTH CONGRUENCY FACTORS AND FOR SPEED OF TARGET. 83
TABLE 2.4 RESULTS OF TWO-WAY REPEATED-MEASURES ANOVA OF SPEED AND DIRECTION FACTORS ON PERFORMANCE SCORES... 83
TABLE 2.5 RESULTS OF TWO-WAY REPEATED-MEASURES ANOVA OF SPEED AND DIRECTION FACTORS ON PERFORMANCE SCORES... 85
TABLE 3.1 DESCRIPTIVE STATISTICS. 110
TABLE 3.2 LINEAR MIXED-EFFECTS MODEL FOR PERFORMANCE BY CONDITION (COMPARED TO CONTROL) AND IOI VARIANCE, CONTROLLING FOR PARTICIPANT. 112
TABLE 4.1 CORRELATIONS BETWEEN DEPENDENT VARIABLES. 139
TABLE 4.2 REPEATED-MEASURES ANOVAS BETWEEN LEVELS OF TEMPO RATIO FOR EACH DEPENDENT VARIABLE. 142
TABLE 4.3 RESULTS OF THE LINEAR MIXED-EFFECTS MODEL FOR THE EFFECT OF TEMPO RATIO AND TIMBRE CONDITIONS 143
TABLE 6.1 LINEAR MIXED-EFFECTS MODEL OF THE INTERACTION EFFECTS BETWEEN CONDITION (SYNCHRONY VS CONTROL, NON-SYNCHRONY VS CONTROL) AND TRIAL NUMBER ON PERFORMANCE. 221
TABLE 6.2 LIST OF VIDEO STIMULI USED IN STUDY 2. 222
TABLE 6.3 LINEAR MIXED-EFFECTS MODEL PREDICTING BONDING FROM TEMPO RATIO, DIFFICULTY AND PUPIL SIZE..... 233
TABLE 6.4 EXAMPLE VIDEOS FOR EACH CONDITION IN STUDY 3..... 234

Statement of Contributions

As an integrated thesis, chapters 1 through to 4 are intended as standalone articles and are presented here as manuscripts to be submitted. I am the lead author on each of these papers, but they were each a collaborative effort. The specific contributions of each author are outlined below.

Chapter 1 – The role of processing fluency in the synchrony bonding effect

Joshua S. Bamford: conceptualisation, investigation, writing – original draft, writing – review & editing. Bronwyn Tarr: conceptualisation, writing – review & editing, supervision. Emma Cohen: conceptualisation, resources, writing – review & editing, supervision.

Chapter 2 – Processing fluency for visual synchrony perception

Joshua S. Bamford: conceptualisation, methodology, software, formal analysis, investigation, data curation, writing – original draft, writing – review & editing, visualisation, project administration. Bronwyn Tarr: conceptualisation, writing – review & editing, supervision. Emma Cohen: conceptualisation, resources, writing – review & editing, supervision, funding acquisition.

Chapter 3 – Synchrony reduces cognitive load during a continuous tapping task

Joshua S. Bamford: conceptualisation, methodology, software, formal analysis, investigation, data curation, writing – original draft, writing – review & editing, visualisation, project administration. Bronwyn Tarr: conceptualisation, writing – review & editing, supervision. Lynden Miles: methodology, analysis, resources, writing – review & editing. Emma Cohen: conceptualisation, resources, writing – review & editing, supervision, funding acquisition.

Chapter 4 – Synchrony-bonding effect mediated by perceived difficulty

Joshua S. Bamford: conceptualisation, methodology, software, formal analysis, investigation, data curation, writing – original draft, writing – review & editing, visualisation, project administration. Bronwyn Tarr: conceptualisation, writing – review & editing, supervision. Lynden Miles: methodology, analysis, resources, writing – review & editing. Emma Cohen: conceptualisation, resources, writing – review & editing, supervision, funding acquisition.

0 Introduction

Throughout this thesis I argue that synchrony is easy. That is, 'easy' in a cognitive processing sense, as some stimuli are easier to perceive and process than others, which is inherently rewarding, a phenomenon often termed processing fluency (Reber et al., 2004). By synchrony, I refer to the type of behavioural coordination in which participants perform actions simultaneously using a predictable and repetitive beat (Bernieri & Rosenthal, 1991). Synchronisation of movement appears to be important in many human activities. Music, dance, marching and rowing, amongst others, require tight synchrony between individuals for the action to be successful (McNeill, 1997). Although such coordination may be ostensibly for specific purposes within the context of the activity itself – e.g., rowing a boat for locomotion – these activities also generate group cohesion and social bonds between participants.

Questions of how people form social bonds have never been more pertinent. Countries such as the UK have begun to recognise a rise in people experiencing loneliness (UK Department for Digital, Culture, 2018). This has only been amplified since the start of the COVID-19 pandemic, which has brought with it a need for social distancing and increased reliance upon disembodied, non-synchronised interactions online (Kumar & Nayar, 2020). Activities which require synchrony, such as music and dance, are difficult to enact in a video call, although many creative solutions have been attempted, and music has played an important part in how some have coped with the pandemic (Carlson et al., 2021; Fraser et al., 2021; Hansen et al., 2021; Krause et al., 2021). Furthermore, there is a perception that synchronised behaviour has been in decline in contemporary Western culture (Garfinkel, 2010; Nietzsche, 1908). However, synchrony may be a vital tool for the formation and maintenance of social bonds.

It has been observed that synchronised activities, such as music and dance, regularly lead to a social bonding effect, leading to suggestions that the evolutionary function of these

activities is primarily for social bonding (Dunbar, 2012b; Savage et al., 2021). Musical behaviour requires a variety of abilities and cognitive processes (Fitch, 2015)(Fitch, 2015). One of the key elements involved in music is an ability to synchronise action (Keller et al., 2014), and the literature surveyed below suggests that this behavioural synchrony is an active ingredient in social bonding effects. Although numerous theories have been suggested, it remains to be explained exactly how synchrony achieves this social bonding effect. This thesis proposes and tests a novel mechanism: that synchrony is easier to process than non-synchrony and is thus subject to processing fluency effects that may lead to increased liking for synchronised stimuli.

0.1 The synchrony-bonding effect

Research focusing on the social outcomes of interpersonal synchrony has described what has become known as the synchrony-bonding effect. This effect has been studied in a wide variety of contexts, from controlled, lab-based experiments to comparisons between real-world activities. In a tightly controlled setting, Hove and Risen (2009) had an experimenter tap either in time with a participant, or at a different tempo, and found that participants rated the experimenter as more likable if they had tapped together in time. Launay, Dean and Bailes (2013) found the effect occurs with both auditory and visual stimuli, and that it can apply to virtual partners (N.B. In this case, the social bonding measure asked about trust rather than likability). The synchrony-bonding effect has been observed in infants, suggesting that synchrony is an important social cue from a young age (Cirelli et al., 2014). It has also been demonstrated in studies of older children in both musical (Kirschner & Tomasello, 2009), and non-musical (Tunçgenç & Cohen, 2018) contexts. However, Stupacher and others (2017) suggested that synchronising to music may have a stronger bonding effect than synchronising to a metronome. Synchrony with members of a minimally constructed out-group may even override negative out-group biases (Tunçgenç & Cohen, 2016). However, social context can

influence whether a participant is motivated to synchronise with a potential synchrony partner (Miles, Griffiths, et al., 2010), and can mediate the subsequent bonding effects (Cohen et al., 2014). While many of these studies have used finger tapping as a simple rhythmic task, Wiltermuth and Heath (2009) observed a stronger effect of synchronous singing compared with moving, possibly because they were singing a national anthem which may have additional meaning to the participants. The synchrony-bonding effect has also been observed in more naturalistic settings involving rowing (Cohen et al., 2010), choir singing (Pearce et al., 2015), dancing in controlled silent discos (Bamford et al., 2016; Tarr et al., 2016; Wen et al., 2017; Woolhouse et al., 2016) and moving in a virtual reality setting (Tarr et al., 2018). Meta-analyses of prior synchrony-bonding research have found a small but consistent effect across different settings (Mogan et al., 2017; Rennung & Göritz, 2016; Vicaria & Dickens, 2016). Taken together, these studies provide evidence for a synchrony-bonding effect, however the precise mechanisms remain unclear.

0.1.1 Mechanisms

Theories abound as to exactly how synchronised movement creates an increased sense of affiliation, trust and bondedness. One account suggests that the endorphin reward system, which is engaged during social grooming in other primates (Dunbar, 2010), is similarly activated when humans synchronise with each other (Tarr et al., 2014). Studies have noted that pain thresholds are elevated following synchronised activity (Cohen et al., 2010; Sullivan et al., 2014; Sullivan & Blacker, 2017; Tarr et al., 2015, 2016), which is used as a proxy measure for endorphin levels since they cannot be measured directly without invasive procedures. Furthermore, music-induced movement without another person present appears to have a similar effect on pain thresholds, and it is possible that subtle movements of the head during music listening may activate the same C-tactile peripheral nervous system that is engaged during social grooming (Dunbar et al., 2021). Others have suggested a role for the oxytocin

system (Harvey, 2020), which has an intricate relationship with the dopaminergic and endorphin systems (Le Merrer et al., 2009; Pearce et al., 2017; Shamay-Tsoory & Abu-Akel, 2016). This preliminary research has begun to establish a neurophysiological basis for the synchrony-bonding effect, however most research has focused on cognitive-psychological mechanisms.

Various cognitive mechanisms behind the synchrony-bonding effect have been suggested. The first of these states that synchronised movement attracts attention, as we like to focus on others synchronised with ourselves (Cirelli et al., 2014). In general, synchronised stimuli will attract attention (Large & Jones, 1999), which extends to the social domain (Macrae et al., 2008). Dyads will tend to orient their head more towards each other when dancing in time in an experimentally controlled silent disco (Bamford, 2017). Since attention is related to memory, the memory for synchronous others is often improved as well (Miles, Nind, et al., 2010; Woolhouse et al., 2016). However, this does not explain why synchrony catches our attention, nor why it would create a heightened sense of affiliation.

Another possibility is that the achievement of shared goals through coordinated action leads people to trust each other more and desire future cooperation (Reddish et al., 2013). In one example, participants were most likely to cooperate after performing a bi-directional synchronous task, compared with either non-synchronous or uni-directional conditions, suggesting a role for shared intentions in which both take an active role in coordination (Reddish et al., 2013). Launay, Dean and Bailes (2013) conducted a similar study, although it also included a condition in which the goal was to avoid synchronising; the largest bonding effects were found in intentional, bi-directional synchrony in which both participants in a dyad actively adapt to the beat of the other (Launay et al., 2013; Reddish et al., 2013). The condition in which participants had to avoid synchrony resulted in less bonding (Launay et al., 2013), however trying to coordinate to avoid synchrony may be a harder goal to achieve than tapping

in time to a shared beat. As will be discussed later, the difficulty of the task may play an important part in the effects of synchrony. The shared goals account is also not mutually exclusive with a joint attention account, as shared intentions may then direct attention towards the other.

As an extension of the shared goals and joint attention theories, the synchrony-bonding effect could be explained through self-other overlap. In order to successfully coordinate with another, one must maintain an internal representation of the other within oneself (Knoblich & Sebanz, 2006). This may be accompanied by activation in action-perception networks in the brain, used to understand another's actions through mirroring them (Overy & Molnar-Szakacs, 2009). The experience of self-other overlap is often related to the rubber hand illusion, which demonstrates that a concept of the self is malleable (Botvinick & Cohen, 1998). In synchrony-bonding experiments the subjective experience of self-other overlap is often quantified using the Inclusion-of-other-in-the-self (IOS) scale (Aron et al., 1992). However, this account applies best to synchrony in dyads, in which two people perform simultaneous movement, and it is not clear whether it could scale to larger groups (Tarr et al., 2014).

An alternative is that synchrony increases the perception that the other is like oneself. Homophily is a more general theory that people tend to like others who are like themselves, and it applies to a wide range of traits and identity markers (McPherson et al., 2001). Evidence of this in relation to synchrony is both scarce and mixed, with some researchers finding that synchrony increased perceived similarity which mediated helping behaviour (Valdesolo & DeSteno, 2011), while others found no effect (Cohen et al., 2014).

Finally, synchronous activity potentially induces positive shared emotional states in participants, which subsequently make them feel closer to each other, commonly described as collective effervescence (Durkheim, 1995) or emotional synchrony (Wlodarczyk et al., 2020).

There is limited support for this theory, as those who synchronise do not necessarily share a positive emotional state (Wiltermuth & Heath, 2009). Crucially, the shared emotion model does not require precise synchrony to function, though a meta-analysis found that precise synchrony does indeed have an effect (Mogan et al., 2017). In light of these findings, there must be a mechanism through which for behavioural synchrony leads to bonding, independent of shared emotions. Those theories discussed so far require higher-order cognitive processes, and some border on being metaphor rather than mechanism. It is nevertheless possible that they might be reconciled through considering the low-level perceptual processing involved in synchrony.

0.2 Fluency through synchrony

Processing fluency is a general psychological principle which may be applied to explain a range of preference judgements. The principle states than an experience of fluency arises from the relative ease and speed in which a stimulus is processed. The experience of fluency is hedonically marked and associated with the brain's internal reward systems (Winkielman et al., 2003). Human brains are energy intensive, so we are incentivised to seek situations that reduce those costs wherever possible (Leder, 2013). Processing fluency may also be used as a reliable cue in interpreting our surroundings: familiar stimuli will be easier to perceive, since they will already exist our memory, and familiarity is safe as we can draw upon prior experience (Reber et al., 2004; Schwarz, 2012), thus processing fluency may be a reliable heuristic (Reber & Unkelbach, 2010). Processing fluency is regularly studied by measuring reaction times, assuming that responding to a stimulus will take longer if it is more difficult to process. Shortened reaction times are often associated with a range of subjective effects, including reduced effort; increased perception of beauty, truthfulness, correctness, familiarity, fame and value; reduced perceived risk; more positive social evaluation; and positive affect

(Reber et al., 2004; Topolinski, 2013). Processing fluency has been considered in relation to synchronised movement, but remains to be empirically tested.

Wheatley and Sievers (2016) were the first to explore the application of processing fluency to the synchrony-bonding effect.[1] Based upon a mirroring account of action perception (Overy & Molnar-Szakacs, 2009), they present an argument for greater neural efficiency when both production and perception of movement are aligned, resulting in social resonance. Wheatley and Sievers (2016) point out that this occurs naturally between individuals in coordination but can be reverse engineered through entrainment to a common musical beat. This would artificially increase the sense of affiliation between individuals, as the other becomes an easy stimulus to process, activating the rewards associated with processing fluency (Winkielman et al., 2003). This provides a new interpretation of the synchrony-bonding effect which is worth investigating.

This thesis aims to test the theory that synchronised movement (as exemplified in music and dance) is easier to process than non-synchronised movement, triggering reward systems that may then be associated with other people present, leading to a social bonding effect. Crucially, this theory should be independent of sensory modality, as processing fluency effects are usually domain general, and also should not require explicitly social stimuli (Reber et al., 2004). The experience of synchrony can be broken down into both perception and participation, as the latter may demand more cognitive effort because it necessarily requires the production of actions that match perception (Keller et al., 2014). With these considerations in mind this theory may be divided up and tested in stages, described below.

[1] Hove and Risen (2009) briefly suggest processing fluency as a possible mechanism in a footnote, but dismiss it because their participants did not report any difference in perceived difficulty of the task. However, without any discussion of fluency as a serious theoretical possibility, and on the basis of only a self-report measure with limited sensitivity, their outright dismissal was perhaps premature.

0.2.1 Chapter 1 - Theory

The first chapter further develops the 'fluency through synchrony' theory. It begins by outlining the evidence for the synchrony-bonding effect, then surveys the existing literature on processing fluency. Processing fluency has been applied to a wide variety of settings, particularly explaining preference judgements for art and music (Reber et al., 2004), but also for social judgements (Lick & Johnson, 2015). The experience of processing fluency is usually induced because of context, stimulus features, or prior experience with a stimulus (Leder, 2013). The standard manipulations used in previous fluency studies, and their relevance to synchronised stimuli, are discussed. Furthermore, the experience of processing fluency usually results in positive affect and increased liking for a stimulus (Reber et al., 1998), which bears similarity to the increased liking and engagement of reward systems that may arise from the perception of synchrony (Hove & Risen, 2009; Tarr et al., 2014). With this background on both the synchrony-bonding and processing fluency effects, the chapter then outlines how synchrony may fulfil the criteria for a fluently processed stimulus.

Perceiving and maintaining synchrony requires the integration of attention, anticipation and adaptation (Keller et al., 2014). Fluency could arise at any of these stages. Synchronised stimuli may be more easily grouped together (Rideaux et al., 2016), resulting in fewer objects to attend to. Synchronised stimuli are also usually more predictable, as they usually follow a regular periodic pattern or rhythmic structure (Vesper et al., 2011), making it easier to anticipate their future actions. When two agents are mutually adapting to each other, this reduces the load on any one agent to adapt (Konvalinka et al., 2010). Predictive processing (Clark, 2015) and Gestalt grouping principles (Wertheimer, 1938) are discussed in relation to these mechanisms, and their possible role in understanding why synchrony could be easier to process than non-synchrony.

Finally, some avenues for further development of the 'fluency through synchrony' account are outlined. The idea that fluency is always preferred is overly simplistic, as there appear to be competing drives for fluency and novelty in aesthetic stimuli (Chmiel & Schubert, 2017; North & Hargreaves, 1995). Understanding synchrony as a source of processing fluency then enables a new research framework to investigate the optimum amount of synchrony to produce bonding effects. Future work could also investigate how the 'fluency through synchrony' theory interacts with other proposed mechanisms for the synchrony-bonding effect. The theory presented here operates at one level of explanation, considering the low-level processing of synchronised stimuli. However, as mentioned above, there are likely higher-order factors contributing to interpersonal coordination, such as shared goals or shared knowledge (Keller et al., 2014). The theory presented here can complement existing theories purporting to explain the synchrony bonding effect. The following empirical chapters report a series of studies that test clear predictions that may arise from this account: that synchrony is easier to perceive than non-synchrony, that synchrony is easier to participate in than non-synchrony, and that the ease of processing that arises from synchrony increases a sense of liking and affiliation.

0.2.2 Chapter 2 – 'Ease of perception in synchrony' (Study 1)

If processing fluency underlies the synchrony-bonding effect, then there is an initial and fundamental assumption that synchronised stimuli are easier to process than non-synchronised stimuli. Study 1 aimed to test this assumption in a non-social visual perception task, using a simple experimental design. Previous studies on processing fluency have observed decreased reaction times for stimuli that are familiar, simple in design, or well defined in terms of figure-ground contrast (Topolinski, 2013). Most of the visual stimuli in previous studies have been static, limiting the inferences that can be made to dynamic and repetitive stimuli. Animated stimuli can have the unfortunate side-effect of influencing reaction times, as

participants may try to entrain their movement to the stimulus (Killingly et al., 2021). Thus, we designed a novel task using stimuli that were moving, but without having a beat or tactus with which to entrain.

The task presented participants with animated Gabor patches (Rideaux et al., 2016), arranged into a flanker formation (Eriksen & Eriksen, 1974), in which the central patch may either be moving in the same or different direction to the patches either side of it. This was similar in design to a study by Lange-Malecki and Treue (2012), except that the Gabor patches allowed for graded, continuous movement. In Study 1a, participants had to respond by indicating the direction of the central Gabor patch, while ignoring the flanker stimuli over a series of trials. A small group of participants were noted to have decreased accuracy on the task when the stimuli were incongruent, but otherwise there was little effect of the manipulation.

In Study 1b, participants completed the same task except that both speed and direction were manipulated. An interaction effect was observed, in which task performance was worst when the stimuli were moving at different speeds, but particularly when the flankers were moving faster than the target. This indicates an effect of non-synchrony on the difficulty of the task, but particularly when the target is less salient than the flankers.

0.2.3 Chapter 3 – 'Ease of participation in synchrony' (Study 2)

Having tested whether synchrony may be easier to perceive under certain circumstances, Study 2 aimed to test whether this holds true when participating in synchrony. It required a different design, based upon previous joint finger tapping studies (Hove & Risen, 2009). In this case, participants kept time by tapping their hand on the desk, in time with one of the two people tapping either in or out of time with each other, side-by-side in a video. A control condition was also included in which only the target person was tapping. This study used a dual-task paradigm, including a secondary task to measure the difficulty of the first,

assuming there would be a trade-off in performance (Huang & Mercer, 2001; Pecenka et al., 2013). Thus, participants also completed multiple trials of a visual search task (Treisman, 1977). This task was chosen as each trial is expected to take longer than some other tasks (e.g., the Flanker task), and certainly longer than the period of the beat in the tapping task (0.5s). If the time taken to complete one trial were less than the period of the beat then participants may wait and respond in time with the beat. Prior research suggests that participants who are provided with a beat are likely to respond to a secondary task in time with that beat (Killingly et al., 2021), which would confound any reaction time measures. Across the three conditions (synchronised, non-synchronised, control), performance on the non-synchronous condition was significantly worse than the other two. Additionally, pupil activity was recorded throughout as a secondary measure of processing load (Beatty, 1982). Effects of condition were found on mean pupil size and number of fixations, however there was no effect on the Index of Pupil Activity (Duchowski et al., 2018, 2020), a novel measure of cognitive load being tested in this study.

0.2.4 Chapter 4 – 'Synchrony-bonding through fluency' (Study 3)

The previous studies provide evidence that both perceiving and participating in synchrony, in either non-social or minimally social settings, is cognitively easier than perceiving or participating in non-synchronous conditions. Study 3 aimed to test whether ease of processing may lead to social bonding. This involved a similar design to Study 2, except without any secondary task, and with the participants and video stimuli using drums rather than tapping on a desk. Processing load was measured using pupillometry, as with Study 2, and a self-report measure. It used a mixed experimental design with a within-subjects factor of tempo ratio (synchrony [120:120 BPM], complementary [120:80 BPM], non-synchrony [120:113.5 BPM]) between the target and distractor people in the video, as well as a between-subjects factor of timbre (matched drums, unmatched drums). Participants spent 60s drumming in each

condition and were then presented with a battery of questions including how much they like the target person, how connected they felt to them, a pictorial measure of self-other overlap (Aron et al., 1992), and how difficult they found the task. We found that subjective measures of social bonding were higher in the synchrony condition. Processing load, measured by self-reported task difficulty and pupil activity, was lower when compared with the non-synchrony condition. No effect was found of timbre. A mediation analysis found that the subjective rating of difficulty mediated the effect of synchrony on bonding. This suggests that processing fluency may explain part of the synchrony-bonding effect.

0.2.5 Chapter 5 – General discussion

The final chapter discusses the theoretical implications of these empirical studies on the evolution of musicality. It is possible, for instance, that early humans were able to leverage quite simple perceptual mechanisms as a means for creating social cohesion. By synchronising one's action with another, one can reduce the other's processing load. This reduced processing load activates internal reward systems (Winkielman et al., 2003), possibly the same endocrine reward systems activated during synchronised movement (Tarr et al., 2014). Individuals with better sensorimotor synchronisation abilities may then be able to gain more benefit from these bonding effects. This should then be discussed from a multilevel selection perspective (Okasha, 2006). Furthermore, the precise perceptual mechanisms through which synchrony is easier to process should be elaborated on. Synchronised action requires a range of cognitive processes (Keller et al., 2014). One possibility is that the subjective experience of self-other overlap, reported by many participants in synchrony-bonding studies (Tarr et al., 2014), may be a perceptual reality as the perceptual streams from both the self and other are grouped together through either Gestalt or auditory streaming principles (Bregman & Campbell, 1971; Rideaux et al., 2016; Shamma et al., 2011). Fewer streams to process may then lead to a reduced processing load. These stimuli may even be streamed together because one stream can be

predicted by another (Bendixen, 2014). Drawing upon embodied predictive processing and the Free Energy Principle (Clark, 2015; Friston & Kiebel, 2009), synchronised, rhythmic actions may be easier to process because they have both a predictable temporal structure (Vuoskoski & Reynolds, 2019), but also because they match one's own actions leading to a shared generative model. Synchrony may spontaneously emerge in mutually inferential systems while agents are attempting to predict each other's actions (Koban et al., 2019; Palacios et al., 2019). Mutual inference leading to synchrony could occur both at multiple biological and time scales. Music and dance then become tools for social 'epistemic foraging' (Clark, 2018), as agents align themselves with each other through mutual prediction (Veissière et al., 2019). It may also be possible to expand this model to include synchrony at different timescales: from harmony to the timing of rituals.

0.3 Summary

This thesis tests a novel theory: that synchronised action induces an experience of processing fluency leading to increased feelings of affiliation and bonding. It does so through three empirical studies, examining whether synchrony is easier to perceive, participate in, and ultimately whether it is associated with increased bonding. Processing fluency is measured throughout using reaction times and pupillometry. Previous theories attempting to explain the synchrony-bonding effect have often relied upon higher-order cognition and lack precise, mechanistic explanations (Cirelli et al., 2014; Demos et al., 2012; McPherson et al., 2001; Mogan et al., 2017; Reddish et al., 2013). However, this new perspective of synchrony as a perceptual processing experience may reveal mechanisms, rather than metaphors, underpinning at least some of the social effects of synchronised action. The notion that synchrony is easy to process may also be consistent with recent theories in cognitive science derived from the Free-Energy Principle (Veissière et al., 2019). By reducing processing load for another through matching their actions, one can effectively manipulate another's internal

reward systems, leading to a social bonding effect (Tarr et al., 2014; Winkielman et al., 2003). Understanding synchrony in this way may change the way we view the role of synchronised action as one of the foundations of social behaviour and culture. Music and dance become valuable tools to early humans, enabling them to bond in increasingly large numbers (Dunbar, 2012b), by leveraging each other's predictive processing (Clark, 2015): a dance of evolution, set in motion because it was easy to process.

1 The role of processing fluency in the synchrony-bonding effect

Joshua S. Bamford[1], Bronwyn Tarr[1,2], Emma Cohen[1]

[1]University of Oxford

[2]The London Interdisciplinary School

1.1 Introduction

Through this paper we will advance the novel theory that the synchrony-bonding effect arises, at least in part, from the experience of processing fluency. Behavioural synchrony has been noted across many cultures and, although there is considerable variation in behaviour, a capacity for synchrony is considered a human universal (McNeill, 1997; Polak et al., 2018). Some non-human animals synchronise their actions for a variety of purposes (Bouwer et al., 2021; Ravignani et al., 2014), but synchrony between humans – particularly in music and dance contexts – appears to have unique social functions (Mehr et al., 2021; Savage et al., 2021). Studies have observed a social bonding effect associated with synchronised action, whereby participants will be more cooperative, pro-social and feel a stronger affiliation towards each other after synchronising their actions (Mogan et al., 2017; Rennung & Göritz, 2016; Vicaria & Dickens, 2016). However, the exact process through which the synchrony-bonding effect is achieved remains unclear. One suggestion is that synchrony blurs the distinction between self and other through the release of endorphins (Tarr et al., 2014), although it is not clear why synchrony should lead to an endorphin response. Based upon general principles in perception and neuroaesthetics (Reber et al., 2004), we suggest that the synchrony-bonding effect arises due to processing fluency: synchrony is easier to process than non-synchrony, leading to a reward response and the release of endorphins. Here we review the existing literature on processing fluency, outlining where it potentially applies to the perceptual experience of synchronised action. We then discuss some perceptual processes that may contribute to the experience of fluency, namely auditory streaming and predictive processing. From this

perspective, we suggest that the experience of self-other overlap is a literal consequence of these processes, and that many of the effects of synchrony are partially attributable to ease of processing.

1.2 Synchrony

Synchrony is a form of interpersonal coordination or joint action, vital for many activities including music, dance, sports, marching and cooperative work (McNeill, 1997). It is closely related to entrainment, which is commonly defined as the process of adjusting behaviour to achieve synchrony (Bernieri et al., 1988), although definitions vary widely throughout the literature (e.g. Lakatos et al., 2019). Bernieri and Rosenthal (1991) divide coordination into two main types: interactional synchrony and behaviour matching. While synchrony primarily describes the timing of actions, behaviour matching (or imitation) describes the kind of action being performed. However, these are not mutually exclusive, and an interacting pair can perform the same action simultaneously in synchrony. Indeed, most of the studies reviewed in this paper involve behavioural synchrony which is matched in both time and kind of action, unless otherwise specified.

Most studies of synchrony have principally involved isochrony, in which actions are performed at exactly the same time, with a regular period (Merker et al., 2009). However, many naturalistic interactions are more complex, involving synchronisation of action at some level, but with a more complex relationship at others (Avitabile et al., 2016). For instance, the drummer in a band may rarely (if ever) be in perfect isochrony with the singer. The hypothetical drummer may play on every beat of the bar, while the singer may perform notes of varying length, with the onset of notes only occasionally coinciding with the drummer. Nevertheless, in this case of simple polyphony they will still perform together ‘in time’ and could be considered synchronised at some metrical level.

For the purposes of this paper, synchrony requires coincidence of timing. A predictable rhythmic structure may help to achieve this, as an individual must be anticipating when an event will occur in order to synchronise with it (Keller, 2013). Neither metrical structure nor isochrony are strictly necessary for synchrony to be achieved. For instance, people can chant or speak in time, however in this case the rhythm of the text provides a cue to help individuals predict the timing of their coordination partners. Synchrony, as time-locked coordination, relies upon temporal prediction more than other forms of coordination.

Further subdivisions of coordination consider the intentionality and directionality of synchrony. Synchrony may be spontaneous and involuntary, or intentional. It may also be unidirectional, in which one agent follows another, or bidirectional, in which both agents are mutually interacting. Bieńkiewicz and others (2021) provide a useful review of some of these factors. The empirical research on the synchrony-bonding effect has mostly used highly controlled settings, and most studies have operationalised synchrony and non-synchrony in their own way.

1.2.1 The synchrony-bonding effect

The synchrony-bonding effect has been observed in a wide variety of experimental settings involving behavioural synchrony. Hove and Risen (2009) demonstrated this experimentally, finding a positive effect of interpersonal synchrony in a finger tapping task on the perceived likability of the tapping partner. The effect can occur with both auditory and visual stimuli (Miles et al., 2009) and for virtual partners (Launay et al., 2013). Participants are more likely to cooperate with each other after engaging in synchronised action vs. non-synchronised action (Reddish et al., 2013), with increased helping behaviour also observed in children after synchrony (Cirelli, 2018). By the age of 12 months, synchrony begins to guide a child's preferences for social (but not non-social) stimuli (Tunçgenç et al., 2015). Other studies have used 'silent disco' paradigms to experimentally manipulate synchrony between

individuals (Bamford et al., 2016; Tarr et al., 2016; Woolhouse et al., 2016), and the use of immersive virtual reality has enabled researchers to isolate the bonding effects of synchrony from common real-world confounds, such as joint intention (Tarr et al., 2018). Meta-analyses suggest a moderate effect of synchrony on prosocial behaviour and perceived social bonding (Mogan et al., 2017; Rennung & Göritz, 2016; Vicaria & Dickens, 2016). However, although people seem to like being in synchrony with each other, and like each other more as a result, the precise mechanisms explaining this effect remain poorly understood.

A range of theories attempt to explain the synchrony-bonding effect. Accounts include the idea that synchrony draws attention to a partner (Bamford et al., 2016; Cirelli et al., 2014); increases one's memory for a partner (Woolhouse et al., 2016); blurs the distinction between self and others' actions in memory (Miles, Nind, et al., 2010); provides evidence of cooperation, which encourages future cooperation (Reddish et al., 2013); creates internal representations of the partner (Knoblich & Sebanz, 2006); increases perception of self-other similarity (Valdesolo & DeSteno, 2011); or evokes shared positive emotion (Demos et al., 2012; Durkheim, 1995). Meanwhile, physiological measurements have suggested a rewarding boost of endorphins and oxytocin following synchronised activity (Cohen et al., 2010; Harvey, 2020; Tarr et al., 2015) which contributes to a blurring of the distinction between self and other (Tarr et al., 2014). Each of these posit that – through either attention, memory, cooperation, internal representation, perceived similarity, shared emotion or endocrine activity – a positive association is formed with another, leading to feelings of bonding, trust, a perceived self-other overlap, and a desire to cooperate.

These and other theories (e.g., muscular bonding, collective effervescence; (McNeill, 1997) attempt to explain the effects of perceiving and acting in synchrony with another. However, there is little detailed consideration of how perceptual processing may contribute to the experience and effects of synchrony. The processes involved in perception are often divided

into selection, organisation, and interpretation of the stimulus (Myers, 2007, p. 237). Synchrony may induce an attentional bias, increasing selection for synchronised stimuli, as prior research has found participants to look towards synchronous stimuli more than non-synchronous stimuli (Bamford et al., 2016; Cirelli et al., 2014; Woolhouse et al., 2016). It remains yet to be explained why our attention should be drawn to synchrony and how synchrony may be involved at other stages of stimulus perception. This is where processing fluency could have some explanatory power.

1.3 Processing fluency

Processing fluency is a general psychological principle which has been applied to explain a range of preference judgements. This principle states that an experience of fluency arises from the relative ease and speed in which a stimulus is processed. It also states that the experience of fluency is hedonically marked, in that fluency is perceived as a positive experience which activates internal reward systems (Winkielman et al., 2003). The processing fluency principle suggests that the brain is incentivised to reduce processing load, and thus prefers stimuli that are easy to process; this is advantageous given the huge energy costs of a human brain (Leder, 2013). In this way, fluency is a kind of meta-cognition in which subjective feeling is used as a source of information about a stimulus (Schwarz, 2012). The hedonic response to fluency may have evolved as a means of determining familiarity for a stimulus, under the assumption that familiar situations are safe, as the individual has prior experience in dealing with them (Reber et al., 2004). Prior exposure to a stimulus makes subsequent presentations easier to process, therefore a stimulus which is easy to process has a higher probability of having been perceived before, relative to a stimulus which is less fluently processed. This reasoning appears almost circular but, as will be discussed, may still serve as a useful heuristic which allows one to assess the familiarity of a stimulus without having to rely upon memory for it. In some cases, however, processing fluency can lead to an illusion of

familiarity, when it arises not because of prior exposure but due to other factors such as features of the stimulus itself. Processing fluency has been used to explain a range of phenomena surrounding preferences for stimuli, and may similarly apply when examining the range of outcomes associated with synchrony.

Processing fluency is observed in the speed and ease in which a task (such as perceiving and responding to a stimulus) is completed. The simplest experiments measure reaction time and accuracy of response as proxies for processing fluency. A range of more subjective effects are also measured, including: reduced effort; increased perceived beauty, truthfulness, correctness, familiarity, fame and value; reduced perceived risk; more positive social evaluation; positive affect; and reduced reliance upon heuristics (Macrae et al., 1993; Reber et al., 2004; Topolinski, 2013). Studies have found independent effects of emotional content and reading speed (in which speed was directly manipulated using scrolling text) on the mood induced by a text, suggesting that speed is its own reward and may lead to the halo of positive associations that fluency possesses (Pronin & Wegner, 2006). Processing fluency may also underly some heuristic decision making – the quick decisions made without full access to all information, and processed mostly subconsciously – because fluency suggests familiarity, and familiarity with a stimulus may be a reliable cue that the stimulus is in some way beneficial to the perceiver (Reber & Unkelbach, 2010). The wide array of effects associated with processing fluency has likely developed because the experience of fluency arising from a stimulus provides useful information about that stimulus when little information is otherwise available.

Processing fluency has become particularly important to the field of neuroaesthetics (Reber et al., 1998), as it provides a means of explaining preferences and responses to both visual art and music (Reber et al., 2004). However, fluency has also been applied to interpersonal judgements (Lick & Johnson, 2015). Fluency may be experienced in response to social stimuli, which could underpin some forms of prejudice towards others who are

unfamiliar or otherwise more difficult to understand (Lick & Johnson, 2015, 2016). Given that interpersonal synchrony often arises in artistic contexts, such as through participation in music or dance, both the aesthetic and social ramifications of fluency may be relevant.

Before investigating the relevance of processing fluency to the synchrony-bonding effect, it is worth examining the ways in which processing fluency has been studied in other domains. After examining how fluency has been applied in other domains, an account of fluency for the processes involved in synchrony can be developed.

1.3.1 Fluency factors

Many stages of processing may exist, depending on the task at hand. Hence there is a need for specificity when considering processing fluency. It is inadequate to state that one stimulus is processed more fluently than another; it inevitably leads one to ask which of the many component processes gives rise to the experience of fluency in this particular case (Topolinski, 2013).

For any given situation, the component cognitive processes, and their potential to produce a fluency experience, must be considered. For example, Leder's (2013) model of aesthetic perception breaks down the processing stages involved in the appraisal of a work of art (Figure 1.1). Initially, Leder (2013) placed processing fluency as a factor only at the implicit processing level, when assessing familiarity with a stimulus, but later acknowledged that fluency may be operating at all levels in the model. Certainly, some of the factors in this model bear a remarkable similarity to the factors outlined as being related to processing fluency by Reber and others (2004). Of particular interest are the 'automatic' processes identified in this model, and their associated features.

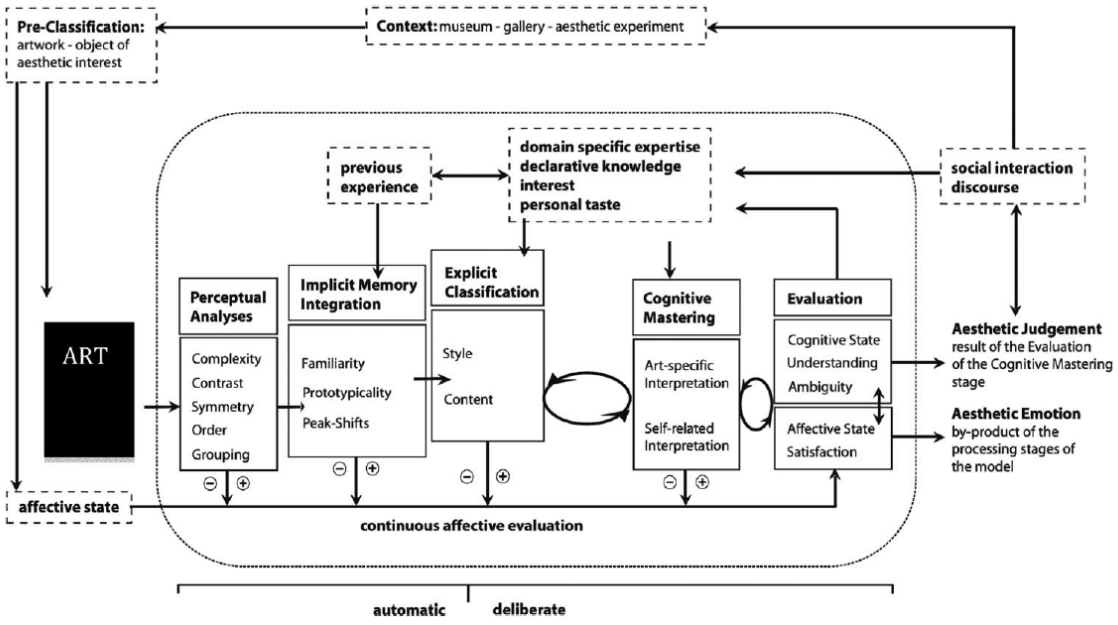


Figure 1.1 Information processing model leading to aesthetic appraisal for a work of art (Leder, 2013).

Drawing from this model, there are several different factors which could lead to an experience of processing fluency. These may involve the context in which a stimulus is perceived, the features of the stimulus itself (complexity, contrast and symmetry), and/or the prior experience the perceiver has with the stimulus (familiarity and prototypicality). Although previous work has treated these as different factors, there may be significant overlap across them. It also must be emphasised that the experience of fluency is not generated by any feature of the stimulus directly, but by the interaction between features of the stimulus and the processing of those features by the perceiver. The three factors already mentioned are primarily relevant to predictions about, and perceptions of, a stimulus. However, a fourth factor relating to production of information or action may also exist. This fourth factor would be particularly relevant to any discussion of processing fluency during synchronous action, as synchrony requires movement production as well as perception.

1.3.2 Context

Stimuli are rarely perceived in isolation, but occur in a greater context filled with other stimuli, and coloured by the perceiver's expectations. Even if fluency does not occur because of the context, the context may influence how fluency is perceived. Context may influence the processing goals of the perceiver (to identify or to understand the stimulus, if these are separable), whether the perceiver has a point of comparison and by what measure the perceiver aims to judge the stimulus.

Studies in neuroaesthetics have suggested that top-down processes change the viewer's perception. Brain regions associated with emotion have been shown to be more active when the viewer expects an aesthetic experience, such as when viewing a work of art, compared to any other visual stimulus which is not framed as an artwork (Cupchik et al., 2009). In this way, the context in which a stimulus is perceived will change the expectations a viewer has of a stimulus, and thus how the viewer will perceive it.

Prior expectations may shape the interpretation of processing fluency. As previously stated, fluency is usually only relied upon with a lack of more specific information (Reber & Unkelbach, 2010). Prior knowledge may include an expectation of fluency (when one expects to see something familiar or unfamiliar), and thus actual experience of fluency becomes especially relevant when it deviates from this expectation (Whittlesea & Leboe, 2003). Furthermore, fluency may be a more reliable cue when it is used comparatively between stimuli (Forster et al., 2015). Thus, fluency is not an absolute quantity, but is relative to either prior expectations or an alternative option.

With the presence of many stimuli, there is a possibility of misattribution of fluency, if the source of fluency is unclear (Schwarz, 2012). Indeed, many experimental methods rely on this, by manipulating features of a stimulus in subtle ways, such that it appears more familiar

when it may actually have never been seen before. In one example, Willems and Van der Linden (2006) found an interaction between the mere-exposure effect and figure-ground contrast, in that one could be annulled by the other. Preference for a familiar stimulus was reduced if the target stimulus had a lower contrast than a distractor (Willems & van der Linden, 2006), although possibly because it was harder to identify the stimulus. As Schwarz (2012) describes in his feelings-as-information theory, feelings are internal, and difficult to attribute to external stimuli.

These examples demonstrate the importance of context to experiences of fluency and may be equally relevant to fluency experienced in synchrony. Synchrony often occurs in specific contexts, so would be expected and even sought out when dancing, but maybe not when walking in a busy street, and this would influence how any resulting perceptual fluency is interpreted. Similarly, dancing with a known partner would create an expectation of synchrony, and so synchrony would not be surprising, while synchrony when dancing with a novel partner may have a stronger fluency effect if the dance is better than expected. These factors must be considered in research designs involving fluency and synchrony.

1.3.3 Stimulus features

Certain features of a stimulus may predispose it to be processed more easily, regardless of any contextual or prior exposure effects. The following are stimulus features that have led to more fluent processing.

1.3.3.1 Contrast

One of the most important features identified in previous research is figure-ground contrast. In the second part of a comprehensive study by Reber, Winkielman and Schwarz (1998), it was found that a higher contrast between a target and the background – presumably making it easier to identify the target – caused the target (in this case, a circle in greyscale) to

be perceived as prettier. Furthermore, text which has a higher contrast with the background, making it easier to read, is perceived as more truthful (Reber & Schwarz, 1999). Increasing the contrast in an image can also make a face appear both more attractive and more familiar, supporting the familiarity-fluency association (Monin, 2003). However, it is worth noting that some authors have failed to observe a change in reaction time related to contrast, indicating that fluency in this case may be due to ease of processing rather than speed of processing (Tinio & Leder, 2009). Their findings were ultimately inconclusive, as they found a complex interaction between reaction time and response accuracy on their stimulus classification task, indicating a trade-off may have been occurring. In any case, increasing figure-ground contrast appears to be a stable method of improving processing fluency for a stimulus, making it more readily identifiable, which may then be interpreted as being more familiar.

Figure-ground contrast is ultimately a question of how easily a stimulus can be distinguished from background noise, and this could occur in any sensory modality. In terms of synchrony, any signal in synchrony with a target stimulus should not inhibit perception of the target, while a signal out of synchrony may be considered noise.

1.3.3.2 Symmetry

Fluency for processing symmetry appears to have been studied less than for figure-ground contrast. In theory, symmetry increases redundancy and thus reduces processing demands (Leder, 2013). Indeed, symmetry could be considered a kind of repetition within a stimulus (Nunes et al., 2015). One study found that symmetrical patterns are judged more positively, and are faster to process, than random patterns of dots (Makin et al., 2012). Although a preference for symmetry is observed in faces and patterns, this does not apply to the same extent in visual arts (Leder, 2013), indicating an effect of context. More research must be done to determine whether there is a specific fluency for processing symmetry, indicating some unique pathways for perceiving symmetrical stimuli, or whether it is a variation on

another factor such as repetition or complexity. Any rhythmic pattern with a highly repetitive structure, such as an isochronous beat, would provide redundancy in the stimulus, just as a symmetrical visual stimulus.

1.3.3.3 Simplicity/Complexity

A stimulus that is more complex should be more difficult to process, and thus less fluent (Leder, 2013). However, defining complexity is difficult. Some authors have suggested using file compression rates as a means of determining the mathematical complexity of a digital stimulus, although there may be differences between psychological and mathematical complexity (Forsythe et al., 2011). In the realm of human movement, other computational methods exist to determine the complexity of a sequence of movements, relating to the rate of change in acceleration (Burger & Toiviainen, 2013). The concept of complexity may also incorporate symmetry (Nadal et al., 2010), and perhaps other factors such as figural goodness.

Figural goodness refers to the clarity of a symbol and is frequently studied in the legibility of fonts and handwriting. Research has found that handwritten essays are subjectively rated as being of better quality – part of the halo of positive effects associated with processing fluency – if the handwriting is easier to read (Greifeneder et al., 2010). This is a prime example of misattribution of fluency, in that it takes longer to understand text and this experience is mistaken for a lack of clarity in the meaning of the text, rather than of the handwriting. However, complexity appears to depend on context and prior experience, as simplicity is not always preferred (Nadal et al., 2010), which will be expounded upon below.

1.3.4 Prior experience

One of the defining features of processing fluency is a sense of familiarity with the stimulus. This can be elicited by features of the stimulus, as discussed above, but may also result from actual familiarity with the stimulus. Prior experience may interact with context as

previously described in forming expectations about a stimulus, or about a particular coordinated action with a target individual or group.

1.3.4.1 Mere-exposure

The mere-exposure effect was first described by Zajonc (1968), as a theory that simply 'repeated exposure of the individual to a stimulus is a sufficient condition for the enhancement of his attitude toward it.' Nothing about the stimulus needs to change, and it does not matter in what manner the individual perceives the stimulus, as long as the subject can perceive it at least subliminally. In fact, it has since been found that subconscious perception of a stimulus may even produce a stronger mere-exposure effect than conscious exposure (Bornstein & D'Agostino, 1992). The mere-exposure effect is often explained through processing fluency, as familiarity with a stimulus makes it easier to process, as one would already have memory of it. The resultant positive affect is interpreted as liking of the stimulus, along with a halo of other positive attributes (Monin, 2003; Reber et al., 2004). As with other fluency effects, the experience of fluency due to mere-exposure provides useful information when information is limited.

1.3.4.2 Duration of exposure

In what may be considered a variation of the mere-exposure effect, changing the duration of exposure may also have similar effects. As part of a battery of experiments testing different processing fluency related effects, Reber and others (1998) found that increasing the length of stimulus exposure also increased subsequent liking for the stimulus (Experiment 3). Winkielman and Cacioppo (2001) similarly found that extended exposure time improved liking for a stimulus, and was also related to a stronger response in facial muscles, recorded through EMG, indicating positive affect. It is, however, worth noting that there may be an upper limit to increasing exposure duration, just as with mere-exposure, as too much exposure may raise the possibility of boredom (Bornstein & D'agostino, 1992). Nevertheless, familiarity through

extended time with a stimulus appears to have a similar effect to repeated brief exposures. This may be pertinent to synchronised behaviours, which often involve an extended period of time attending to a partner or a group (Bispham, 2018), and thus the salience of those others may increase simply through consistent exposure.

1.3.4.3 Perceptual priming

Another related concept is that of priming, which involves partial exposure to a stimulus, or features of a stimulus. A stimulus which has been preceded by a matched prime (as opposed to an unmatched distractor) is perceived as nicer to look at and quicker to recognise (Reber et al., 1998), as well as eliciting a facial response indicating positive affect (Winkielman & Cacioppo, 2001). Debates in the literature over the relationship between priming and mere-exposure remain largely inconclusive (Butler & Berry, 2004). Both likely rely upon simple perceptual processes and the results appear to be quite similar. In a rhythmic stimulus, each beat may prime the next, assuming the temporal pattern is maintained, although it would be difficult to disentangle this from other effects arising from prior experience.

1.3.4.4 Conceptual priming

Whereas perceptual priming (and mere-exposure) aids in stimulus recognition by using features of the stimulus as a prime, conceptual priming uses the meaning of the stimulus (Butler & Berry, 2004). One example of this is priming a painting using either a title related to the content of the painting, which increases appreciation for the painting, as compared to an irrelevant title (Belke et al., 2010). Similarly, the perceived truth of a statement is enhanced through repeated exposure, regardless of whether the statement is paraphrased (Silva et al., 2017). In a study by North, Hargreaves and McKendrick (1999), it was found that shoppers would prefer to buy wine that matched the background music by country of origin. In all of these cases, preferences were shaped because of the meaning or concept behind a stimulus, rather than the specific feature. This suggests cross-modal priming is possible, which

presumably relies upon higher-order cognitive processes. Conceptual priming may be less relevant in the perceptual experience of participating in synchronised action, but may contribute in the social context around synchrony. For instance, talking about synchronising prior to engaging in a synchronised activity could enhance an experience of fluency.

1.3.4.5 Motor-priming

Another special case of cross-modal priming is based on embodied cognition principles and is particularly relevant to discussions of dancing. It relies upon an assumption that motor processing is important to understanding some stimuli. This approach was initially tested on the mere-exposure effect and found that instructing participants to engage with a motor distractor task (chewing gum) eliminated the mere-exposure effect for reading words (Topolinski & Strack, 2009). This has since been applied to ocular muscles, in an experiment where subjects completed eye movements as either a matched or unmatched prime for a moving image. When the ocular motor prime was matched, there was higher preference for the perceived movement (Topolinski, 2010). It has also been shown that appreciation for an artwork may be increased if the viewer has previously performed arm movements that simulate the brush strokes used by the artist (Leder et al., 2012).

The concept of motor-priming to induce fluency presents a highly embodied approach to neuroaesthetics – one in which the prior actions of the viewer (rather than their passive viewing experience) increase their experienced fluency for a stimulus. Motor-priming could be particularly relevant to experiences during synchronised action. If two actors are moving in synchrony, then their own actions would prime them to perceive the other's actions.

1.3.4.6 Prototypicality

Aside from repeated exposures and recent priming, fluency may also arise when a stimulus conforms to our expectations for a category of stimuli. In visual perception, this

applies for geometric shapes as well as people, as Lick and Johnson (2013) demonstrated. The authors found that participants experienced processing disfluency (measured by reaction times) when perceiving people who were difficult to categorise (e.g., based on sexual orientation) and that this was a mediating factor in explaining prejudice against homosexual targets (Lick & Johnson, 2013). However, no such effect was found on more overt differences such as race (Lick & Johnson, 2013). This suggests that fluency through prototypicality arises from an ease of classification according to our previously held stereotypes. In essence, it is not fundamentally different to other familiarity effects, except that instead of a familiarity with the exact stimulus (mere-exposure) or a feature of it (priming), prototypicality refers to a familiarity with a category of stimuli. This could occur at larger timescales in some synchronised movement contexts: for instance, in recognising dance moves, if familiar with a particular dance style.

1.3.5 Fluency of Production

Up until this point, processing fluency has been described as a response to perceiving a stimulus, albeit through an interaction with the perceiver's expectations and experience. Another class of effects has also been observed in the production of responses, now termed retrieval fluency.

1.3.5.1 Retrieval fluency

This effect refers to ease of recall for a memory item. The experience of retrieval fluency arises when recalling a memory item is quick and easy, and leads to the belief that the memory is more correct and valid (Kelley & Lindsay, 1993). This effect has demonstrated that people are more confident that they have the correct answer to a question when the answer comes to them faster (Koriat & Levy-Sadot, 2001). It is suggested that this observation has two components: familiarity for the cue, and subsequent accessibility of the item being retrieved from memory (Koriat & Levy-Sadot, 2001). However, the method used in this study may not

have entirely distinguished between these two processes, as they were manipulated by changing the way in which the question was asked.

A further study suggested that the speed at which an answer is found is not limited to recall. Topolinski and Reber (2010) showed that if the solution to a problem is presented to a subject more quickly after the initial problem, it is perceived to be more correct. This may demonstrate a preference for fast answers, whether the answer comes from an internal (memory) or external source (Topolinski, 2013). In this way, the effect may not be fundamentally different to the previously discussed fluency mechanisms.

1.3.5.2 Action execution

Although most studies on processing fluency have involved perceptual tasks, the possibility of fluency for motor tasks should be considered. Certainly, some success has been achieved through investigating motor-priming effects (Leder et al., 2012; Topolinski, 2010). In theory, a kind of retrieval fluency for motor actions (e.g. through recall of a choreographed sequence of movements) could be imagined. Similarly, some actions may be easier to perform for purely physical reasons (e.g. tapping a beat at a moderate pace, compared with one that pushes the limit of muscle control). These questions are yet to be explored, however there are other issues for the processing fluency principle in general which should be addressed.

1.3.6 Competing drives for fluency and novelty

With the evidence up to this point, one could believe that higher processing fluency always leads to the most positive experience and all the heightened liking, perceptions of truth and beauty, and positive affect which that implies. However, there may be upper limits to the fluency effect. In the neuroaesthetic study of art preferences (be it visual arts or music), there appears to be an optimum balance between both novelty and familiarity, and complexity and simplicity. The optimum level of complexity for a work of art may depend to some degree on

the viewer's prior experience (Nadal et al., 2010). This is similarly observed in musical preferences (Chmiel & Schubert, 2017; North & Hargreaves, 1995). Research into individual differences has also associated the Openness personality trait, with preference for novelty in both visual (Fayn et al., 2015) and auditory (Hunter & Schellenberg, 2011) stimuli. There may also be effects of gender and culture on preferences for complexity (Street et al., 2016). The nature of this fluency optimum may be elucidated through comparison to related concepts such as free-energy minimisation and flow states.

The free-energy principle can be seen as an extension of predictive coding. It suggests that all agents are driven to reduce uncertainty (free energy) through creating generative models of their surroundings (Friston & Kiebel, 2009). 'Perception' is then redefined as prediction error, which is used to fine-tune the model using Bayesian principles (Friston & Kiebel, 2009). Crucially, it is more efficient for the agent to be able to rely upon their internal models, rather than having to constantly correct errors arising from new and surprising information, which can be seen in the neural activity for unexpected stimuli (Malmierca et al., 2014). Thus, the free-energy principle could be underpinning some of the processing fluency effects, while also accounting for the observed preference for an optimum level of complexity.

Some would suggest that the drive to reduce uncertainty should necessarily mean that all agents should prefer the simplest possible environment with minimal stimuli: a critique known as the dark room problem (Sun & Firestone, 2020). However, it is not a lack of stimuli that serves as motivation in the free-energy principle, but rather the action of fine-tuning the model, which requires access to new information (Friston et al., 2012). Successfully updating one's model may then provide positive, affective rewards (Van de Cruys, 2017). Under the free-energy principle, the concept of optimum (rather than maximum) fluency begins to make sense.

In this formulation, novel or unusually complex works of art or music likely violate the viewer's expectations. The interplay between expectancy and surprise may be used to maintain aesthetic interest (Koelsch et al., 2019; Van de Cruys & Wagemans, 2011). In some cases, ambiguity may be preferred, allowing the viewer to experience and overcome the challenge of understanding the stimulus (Jakesch et al., 2013). The predictive brain is constantly working to reduce uncertainty, but if there is no uncertainty left to reduce then this drive cannot be fulfilled. Thus, there may be two subconscious driving forces in competition: the need to seek novelty and avoid boredom, and the desire for the safe and familiar.

At a higher timescale, it's possible that achieving the balance between familiarity and complexity leads to a flow state. Flow states have been described when someone is at peak performance and, crucially, when the challenge posed by a task is perfectly matched by their skill at the task (Csikszentmihalyi & Csikszentmihalyi, 1988). This is often studied using subjective self-report measures, in which participants describe a positive experience, being focused solely on the task at hand, often losing track of time (Jackson & Marsh, 1996). It is commonly experienced in music (Chirico et al., 2015), sports (Zumeta et al., 2016), and more generally during interpersonal synchronisation (Stupacher, 2019). Furthermore, it has been suggested that a shared flow state may contribute to the bonding effects of synchronised action (Zumeta et al., 2016).

Taking these theories into consideration may help to clarify the processing fluency effect. The experience of stimuli that conform to the characteristics outlined previously (simplicity, familiarity, etc.) produce positive affective rewards, and are then viewed favourably (Winkielman et al., 2003). Part of the fluency may arise because these stimuli conform to predictions made about the world, and thus produce less error than more complex or unfamiliar stimuli (Friston & Kiebel, 2009). However, this ceases to be a positive experience if stimuli are too predictable, as there is a drive to always improve one's predictive model

(Friston et al., 2012). This could explain why people often seek an optimum level of novelty or complexity in a stimulus (Chmiel & Schubert, 2017; Nadal et al., 2010). When actively performing a task, achieving the optimum balance in task difficulty may then facilitate a flow state (Csikszentmihalyi & Csikszentmihalyi, 1988). Using this theoretical basis, we may then begin to examine how ease of processing could relate to the perception of synchronised stimuli and participation in synchronised action.

1.4 Fluency in synchrony

The principles of processing fluency may apply to the perception of, and participation, in synchrony. Just as stimuli that are easier to perceive are more likable (Winkielman et al., 2003), so a person who is easier to perceive may also be more likable (Lick & Johnson, 2015). Synchronising actions may be one method of reducing processing load in social stimuli, leading to the range of social outcomes associated with the synchrony-bonding effect.

The observed outcomes of the synchrony-bonding effect are not dissimilar to the effects that arise through processing fluency. Synchrony has been observed to enhance self-reported likability of a synchronous target (Hove & Risen, 2009), promote prosocial behaviour (Cirelli, 2018; Kirschner & Tomasello, 2009), improve performance on subsequent cooperation tasks (Lang et al., 2016), increase self-esteem (Lumsden et al., 2014), encourage visual attention towards and physical closeness to the synchronous other (Bamford et al., 2016; Wen et al., 2017), improve memory for synchronous others (Woolhouse et al., 2016), increase willingness to share emotions (Koole et al., 2020), and increase opioidergic activity, as measured through increased pain thresholds (Tarr et al., 2015). Some of the observed synchrony effects appear similar to the observed effects of processing fluency, in particular the improved preference, likability, familiarity, trust, memory recall, and visual attention. The other effects associated with processing fluency, such as perceived simplicity, value, risk and beauty, were not measured in the synchrony-bonding studies reviewed here, but that does not preclude them.

We suggest that some of the effects of synchrony may be explained through processing fluency. Wheatley and Sievers (2016) were the first to explore the application of processing fluency to the synchrony-bonding effect. Based upon a neural mirroring account of action-perception (Overy & Molnar-Szakacs, 2009), they present an argument for greater neural efficiency when both production and perception of movement are aligned in a kind of social resonance. They point out that this occurs naturally between individuals in coordination but can be reverse engineered through entrainment to a common musical beat, hence artificially increasing the sense of affiliation between individuals (Wheatley & Sievers, 2016). This provides a new interpretation of the synchrony-bonding effect which is worth investigating.

The problem with processing fluency as a general psychological principle is that it can become so broad and all-encompassing that it ceases to be useful as an explanatory account of anything in particular. Everything the brain does is processing, and thus processing fluency has been invoked in numerous theories, but specific fluent processes are not always identified. If synchrony is easier to process than non-synchrony, which process is being experienced fluently must be established. For that, a detailed model of the processes involved in synchronisation of movement is required. This could be similar to the model proposed by Leder (2013) for art appreciation. Joint action research provides a very good basis, as many mechanisms involved in perceiving and matching the actions of others have already been described (Sebanz et al., 2006). The model proposed by Keller, Novembre and Hove (2014) may also provide a useful starting place (Figure 1.2). If synchrony is easy to process, and this in turn leads to prosocial consequences due to the hedonic marking of fluent processing, then we must begin by identifying the cognitive processes involved in synchronising.

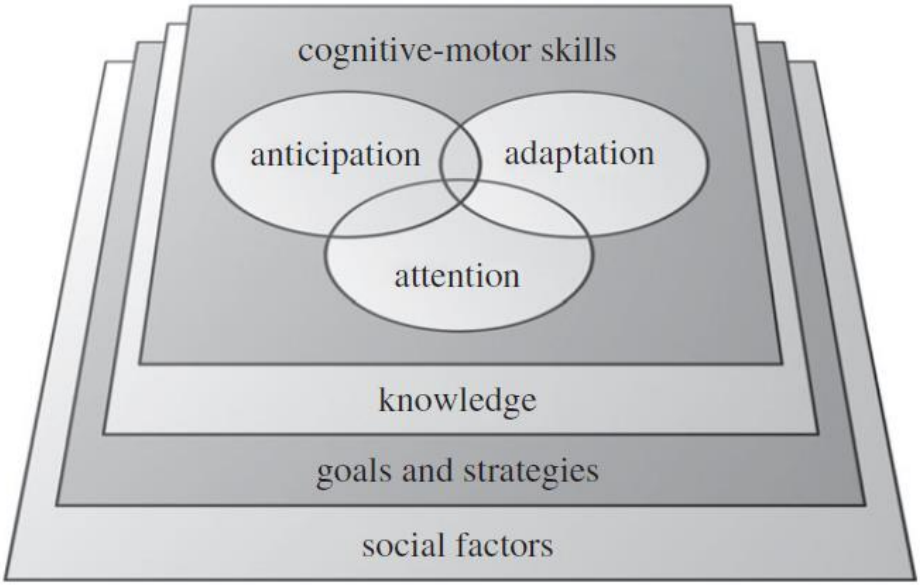


Figure 1.2 Factors that affect interpersonal coordination (Keller et al., 2014).

Keller, Novembre & Hove (2014) propose an account of cognitive mechanisms involved in interpersonal coordination, which may serve as a useful starting point. Three core processes are identified: attention, anticipation, and adaptation. To coordinate movement, perceptual attention must be directed towards the actions of both the target and the self, the target's movement must be anticipated, and one's own movement must be adapted, both to the anticipated and actual movement of the other (Keller et al., 2014). These three processes occur concurrently and in constant interaction with each other. This makes interpersonal coordination a cognitively demanding task, and sufficient cognitive resources must be made available. An increase in the difficulty of any of these three cognitive processes may result in less fluent processing overall, while if the aims of any of these processes are easier to complete, a sense of fluency – with all the associated benefits already discussed – may be experienced. What follows is an examination of the role each of these processes may have in an experience of fluency during synchronised action.

1.4.1 Ease of attention

A stimulus that is synchronous either with other stimuli or with the perceiver themselves may be easier to attend to than a non-synchronous stimulus, simply due to the reduced complexity of the stimuli present. The literature on Auditory Scene Analysis (ASA) suggests that synchrony may result in two stimuli being perceptually grouped together (Bregman, 1990). If one perceptual stream is simpler to process than two, then synchrony may give rise to processing fluency through ease of attention. Furthermore, if one participates in synchrony with another, and aspects of both the perception of self and other are streamed together through ASA, then this would suggest a very literal interpretation of the self-other overlap experienced subjectively during synchrony tasks.

1.4.1.1 Auditory Scene Analysis

Auditory Scene Analysis (ASA) outlines a process of perceptually grouping sounds into streams (or auditory objects), according to their source (Bregman, 1990). It is similar in many ways to Gestalt grouping principles outlined by Wertheimer (1938), although focused specifically on audition. Localisation of sound sources poses specific challenges, due to the nature of acoustics and the auditory sensory system (Pressnitzer et al., 2011). Sound waves from all sources are initially superimposed upon each other at the eardrum. This single audio input is then divided up by the cochlea into multiple channels based on frequency. The brain must then make sense of this sensory input. Correct auditory streaming when there are multiple sound sources present is often referred to as the 'cocktail party problem,' and relies on certain top-down processes (Pressnitzer et al., 2011). Disambiguating multiple overlapping streams is then more effortful than a single stream with an apparent single source.

Sounds are streamed according to two main classes of cues: horizontal and vertical. Vertical organisation of sound relates to pitch and harmony, while horizontal organisation is done according to time (Pressnitzer et al., 2011). One of the strongest factors in ASA is

synchrony of onset, as sounds that occur at the same time often originate from the same source (Hukin & Darwin, 1995). This is an important heuristic, learnt through experience in one's environment.

Alternatively, instead of distinctive horizontal and vertical cues, the auditory system performs a coherence analysis of all frequency channel inputs, and groups them into streams based on similarity, in which temporal coherence is the most important factor (Elhilali et al., 2009). This process may also be mediated by attention, through prioritising some features over others (Shamma et al., 2011). In either case, cognitive difficulty may arise when sounds are not easily streamed.

Common to both visual and auditory grouping/streaming is the phenomenon of bistable illusions. This occurs when the features of stimuli are not easily grouped, and thus can be interpreted in two ways, depending on top-down attentional processes. There are many examples of this in visual perception, such as the vase illusion (Figure *1.3*). A similar phenomenon has been studied in auditory perception, in which alternating high and low tones are played rapidly. Participants could either perceive this pattern as a single melody line or as two separate lines (van Noorden, 1975). This perceptual ambiguity is often impossible to resolve, as the perception of bistable illusions may spontaneously swap between interpretations.

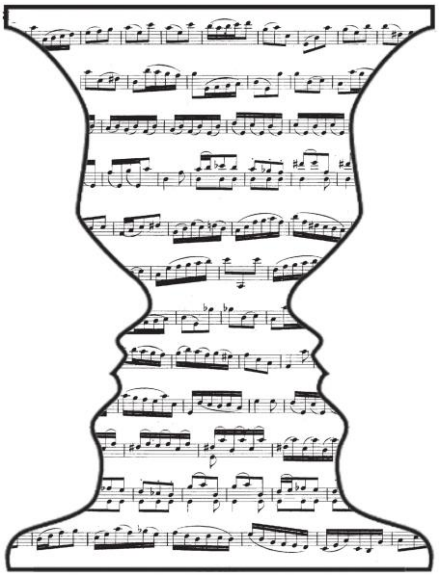


Figure 1.3 Visual bistability, in which the image may be seen either as two faces or one vase (Pressnitzer et al., 2011).

It is possible that ambiguity in ASA contributes to disfluency of processing as one attempts to attend to conflicting auditory streams. The reverse may also be true, that the experience of self-other overlap arises from fluency in auditory streaming. Interestingly, even conflicting auditory location cues may not always negate synchrony cues for streaming (Micheyl et al., 2010). Grouping different individuals into one stream, despite spatial differences, is thus entirely possible. If one taps a drum in exact synchrony with another individual, then these two auditory sources may be processed as one stream. However, slight asynchrony may increase the attentional load as one attempts to resolve the stimuli. Increased attentional load has been shown to increase processing time required to identify tones (Alain & Izenberg, 2003), and ambiguous streams have been associated with greater activity in brain regions associated with attention and executive function (Dolležal et al., 2014), implying less fluent processing. Thus, using an ASA framework we may manipulate fluency for the auditory perception of synchrony.

Processing disfluency may not only arise from ambiguity in auditory streaming. Studies have found that auditory stimuli that are perceived to be from multiple sources require more

cognitive resources, as are those that are more unpredictable, and that these are independent effects (Hughes et al., 2013; Zeamer & Fox Tree, 2013). Synchronous stimuli may be perceived as originating from the same source, and are thus less cognitively taxing to perceive than non-synchronous stimuli.

An analytical possibility may exist in Ellis' (1999) computational model of coherence in auditory streaming, if this could be applied to the sound output of two individuals and used to predict how a listener would stream their sound. Furthermore, given that Gestalt grouping of visual stimuli may also be done on the basis of synchrony (Rideaux et al., 2016), a similar approach could be applied to the visual perception of synchrony. However, the grouping of perceptual objects in this case would likely rely upon a multi-modal process involving proprioception. When synchronised with a target stimulus, the auditory streams between self and other may be indistinguishable, however the visual perception is usually very distinct; subjects may hear themselves tapping but cannot see themselves. Therefore, any synchrony with a visual stimulus would probably be matched with the proprioceptive sensation of movement, requiring cross-modal grouping.

1.4.1.2 Attention resource allocation

ASA principles may also apply when considering the attentional processes of a performer coordinating their actions in an ensemble. This complex process may be referred to as Prioritised Integrative Attending (PIA; Keller, 2001). PIA requires a balance of attentional resources between the self and other, and studies have found that increased rhythmic complexity of a performed action, or reduced coherence between one's performed rhythm and that of the other, reduces the capacity for other-directed attention (Keller, 2001). This theory does not explicitly suggest the possibility of self-other streaming, although self-other streaming would be consistent with the increased capacity for other-directed attention when there is greater rhythmic coherence between self and other.

The process of dividing attention between self and other is not exclusive to musical ensembles. Evidence from animal studies exists for a neurological process which selectively suppresses the processing of auditory feedback from movement related sounds in general (Schneider et al., 2018). This mechanism would allow attentional resources to be focused on relevant environmental sounds, without distractions from the acoustic consequences of one's own movement. Environmental sounds that occur in time with one's own movement related sounds would then face the possibility of being streamed together with sounds from the self and may fall below conscious attention. Not only would this be easier to process, but it may give rise to the feeling that sounds from the other are part of an expanded sense of self.

1.4.2 Ease of anticipation

Moving in synchrony requires going beyond reacting to the actions of another, but actively predicting their actions. Actions that are more regular and predictable are easier for a perceiver to predict. When actions conform to one's predictions, they could be easier to perceive, hence there is a possibility of processing fluency arising from the anticipation phase of the cognitive mechanisms in synchronous action (Keller et al., 2014). The literature on predictive coding, a rapidly developing theory in cognitive science, may be drawn upon to develop this further.

1.4.2.1 Embodied predictive coding

Perception psychology has undergone many major revisions, gradually shifting from an initial perspective that perception was mostly a bottom-up process, with stimulus features feeding up from sensory organs to the brain, to the concept of predictive coding, in which most processing is assumed to be top-down. This theory proposes that stimuli are predicted by the brain, rather than perceived (Rao & Ballard, 1999). The only information to flow through from the senses is prediction error, which is then used to update the internal predictive modelling (Clark, 2015). This also proposes a radical view of attention, in which attention is directed by

prediction error, rather than primarily top-down processes (Clark, 2015). In theory, this method of predictive modelling and error detection is more efficient as long as prediction error – or free-energy – is minimised (Friston & Kiebel, 2009). The processing of error could thus be a source of processing disfluency, as it forces the updating of a stable internal model.

Predictive coding is not limited to perception, but also presents a radical view of motor planning. According to the predictive account, motor processing is no different to visual or auditory processing, except that it deals with the sense of proprioception (Clark, 2015). The visual cortex tests predictions for visual stimuli while the motor cortex predicts proprioceptive feedback from the body. Since the predictive brain aims to reduce prediction error, signals to the muscles then stimulate them to move in such a way that the prediction is confirmed (Clark, 2015). In this way, the motor cortex does not have precise control of the limbs, but simply stipulates a predicted sensory state, and then allows the rest of the nervous system to find the most efficient way of achieving that state.

This approach rejects the bottom-up approach to cognition, in which there is a sensory input, which is processed and then followed by a motor output. In the predictive account, motor action is not a response to an input, but rather a way of selecting the next input, as the higher levels of cognition make predictions and then lower levels seek to confirm those predictions (Clark, 2015). Consistent with a processing fluency account, emotions then provide constant feedback on the effectiveness of our predictions (Van de Cruys, 2017). Thus, the confirmation of predictions may be considered a fluent process, which leads to positive affect.

1.4.2.2 Action prediction in synchrony

The predictive coding theory just described would suggest that an individual is constantly working to predict another's actions when observing them. Indeed, people do seem to maintain an internal representation when observing another's actions, regardless of whether

they are trying to coordinate with the target or not (Obhi & Sebanz, 2011). Correspondingly, an individual may attempt to make themselves more predictable when they know someone else is trying to coordinate with them. This was observed in a study where participants exhibited more predictable timing when asked to synchronise with another, compared to if they were asked to only respond to the other's action (Vesper et al., 2011). Modelling/predicting the other, and enhancing one's own predictability, seem to be integral to coordinated action.

The predictability of those actions may then influence the fluency with which they can be processed. According to Clark (2015), prediction is an embodied process, relying upon internal modelling built up through experience. When one coordinates actions with another, they would then use themselves as a reference point. This is exemplified in Chapman's (1968) 'Outfielder's Problem,' in which it is proposed that someone aiming to catch a ball does so not by modelling the motion of the ball, with the understanding of Newtonian physics that would require, but rather through matching the ball's motion to their own. The ball is thus easier to catch if the outfielder runs with the ball, aiming to keep a consistent relationship with the ball, until their trajectories intersect and the ball can be caught (Chapman, 1968). It is possible that this same process is used to coordinate actions with others. Under this model, the actions of another become more predictable when they are kept in a constant relationship with the self. In synchrony, one may assume that the other will keep moving with the self, thus making prediction of the other partially predicted by movement of the self, making for a more efficient process.

In addition to being easier to process, predictability may also be used as a cue for self-recognition. Across multiple sensory domains, we are able to filter out stimuli that arise from oneself, possibly because they are more predictable, which has been used to explain why most humans cannot tickle themselves and are not distracted by their own footsteps (Brown et al., 2013). Stimuli derived from one's own actions should be predictable from one's own internal

models, and this predictability may be used as a cue to identifying oneself in the environment (Apps & Tsakiris, 2014). If the actions of another become highly predictable, and even match one's own actions, then there could be a possibility of misidentifying them as originating within the self. This could be similar to the idea of self-other streaming posited in the previous section.

1.4.3 Ease of adaptation

Ease of processing in adaptation should also be considered, relating to fluency of action execution discussed previously. In this category, there may be two ways in which adaptation is made easier: firstly, through familiarity with the movement one is synchronising to, in the case of kind-locked synchrony (total synchrony in which individuals perform the same kind of action at the same time); secondly, through mutual adaptation in bidirectional synchrony, reducing the adaptation load on any individual in the system.

1.4.3.1 Motor familiarity

This would relate to kind-locked coordination more than time-locked, as the timing may be less important than the kind of action. Previous studies have used motor-priming techniques to increase familiarity with a movement, which has been shown to lead to subsequent experiences of processing fluency and increased liking for a stimulus (Topolinski, 2010). If this is the case, then repeated exposure to a specific movement may make that movement easier to perform. Conversely, performing a movement may make that movement easier to perceive. As naturalistic dance settings involve long sequences of movements which may be performed repetitively by participants, this sets up a platform for playing with familiarity of movement.

1.4.3.2 Mutual adaptation

Bidirectional synchrony may occur when two (or more) agents attend and adapt to each other. Approaching synchrony as a dynamical system between two oscillators, it should be easier for each oscillator if both share the load of adaptation, as opposed to a situation where

one oscillator must adapt to an unchanging tempo (Fairhurst et al., 2013; Goebl & Guggenberger, 2015; Konvalinka et al., 2010). This could have an interesting interaction with a predictive coding account, as an oscillator which freely adapts to its counterpart would be less intrinsically predictable, yet may be more predictable for its counterpart if both are similarly inclined to adapt to each other. Individual differences have been noted in the adaptation ability (or willingness) of human subjects (Carlson et al., 2020; Mills et al., 2015), suggesting that some people may be better at easing another's processing load during coordination. If mutual adaptation (bidirectional synchrony) is easier than unidirectional synchrony, this may lead to a feeling of processing fluency.

1.4.4 Integrating attention, anticipation and adaptation

While it may be theoretically convenient to separate predictive processing and Auditory Scene Analysis (ASA), there may be some overlap. This remains consistent with the model of Keller and others (2014), in which there is substantial overlap between the processes of attention, anticipation and adaptation, however the extent of the overlap is not known.

As previously mentioned, the predictive coding theory purports to include attentional processes as well. Attention can be viewed as weighting within the internal statistical model, and is heavily directed by prediction error (Clark, 2015). Within ASA, there is a possibility that predictability may serve as a strong cue for auditory streaming (Bendixen, 2014). A review by Bendixen (2014) suggests that the predictability of both a foreground and background stream may cumulatively contribute to the ease of segregating foreground and background into separate streams. A background auditory stimulus may also be more distracting when it is less predictable (Hughes et al., 2013). Unpredictability in auditory streams would seem to generate ambiguity in the auditory scene, increasing cognitive load.

Mutual adaptation may also be explained through the lens of predictability. As noted earlier, people will act to make themselves more predictable when trying to coordinate with another (Vesper et al., 2011). Bidirectional synchrony could be described as a process through which both members of a dyad work to make themselves more predictable.

Keller and colleagues (2014) note that the central mechanisms of attention, adaptation and anticipation reside within a framework of (shared) knowledge, goals and social expectations. The expectation of synchrony, depending on social context, may indeed influence how it is experienced. This may also be viewed through a predictive coding lens, as individuals who share expectations of synchrony, through shared cultural knowledge encoded as predictions of other's actions in a given social situation, may be more likely to coordinate well (Ramstead et al., 2016). Conversely, if one participant violates the others' expectations of synchrony, this may be viewed unfavourably. The most interesting case is perhaps when synchrony between individuals is achieved more easily than predicted, as the greatest processing fluency effects are thought to emerge when processing is not just easier, but easier than expected (Reber et al., 2004). Expectations of synchrony may indeed be shaped by other factors such as group identity (Macpherson, Fay, et al., 2020). Any future research should take care to consider contextual factors that may influence the experience of fluency, and its interpretation.

1.5 Summary

This review has outlined the principles of processing fluency and discussed how they may contribute to the synchrony-bonding effect. Although the pathways involved in interpersonal synchronisation can be divided into attention, anticipation and adaptation, there is significant overlap between these processes built into the model in Figure 1.2. Any research programme aiming to test this novel theory could examine any of these component processes involved in perceiving or maintaining synchrony. Attention would likely be the first step,

however, as if one does not notice the other moving in time with oneself, one cannot bond with them.

Auditory streaming suggests that temporal alignment of sounds may appear to be from the same source are perceptually streamed together (Bregman, 1990), being perceived as one auditory object. When stimuli include one's own actions, this relates to the integrative attention problem outlined by Keller (2001), and the need to distinguish between self and other when prioritising attention to relevant environmental sounds (Schneider et al., 2018). If two objects (be they two others, or self and other) are perceptually streamed and processed as one, then they would demand less from one's attentional capacity leading to more fluent processing.

From a predictive coding perspective, attention is directed by errors in prediction (Friston & Kiebel, 2009). This would suggest that, in joint action, two or more agents are mutually predicting each other's actions (Friston & Frith, 2015a). When the other deviates from one's predictions, one is forced to update one's predictions (Clark, 2015). These predictions may be multilayered, when aware that one's actions could influence the future actions of an interaction partner (Veissière et al., 2019). The best course of action, to optimise synchrony, may be to make oneself maximally predictable, thus easing the processing load for the other (Koban et al., 2019; Konvalinka et al., 2010). Furthermore, predictability of a stimulus may be used as a cue to perceptual grouping (Bendixen, 2014), indicating that auditory streaming could be underpinned by predictive perception.

Although it is tempting to place predictive coding within the anticipation section of Keller and others' (2014) model, there may be substantial overlap between these pathways. Indeed, predictive coding could encompass attention, anticipation (prediction) and adaptation (action), within the active inference framework (Friston & Frith, 2015a). When both are

moving in synchrony, the other must be moving in accordance with one's own predictions, and thus no extra attentional capacity would need to be directed towards them.

This account suggests that the perception of a synchronous body should be easier than the perception of a non-synchronous body. The perceptual consequences of another's movement may be streamed in with our own, and predicted based upon our own actions, requiring fewer attentional resources. Reduced demand for attentional resources, and fulfillment of one's predictions, would be experienced as processing fluency, leading to an affective reward (Winkielman et al., 2003). There may, however, be an optimum amount of fluency, and more recent work finds a balance between predictability and novelty must be struck in order to maintain interest (Clark, 2018). Sustained interaction that maintains this optimum balance could even induce a flow state in participants (Stupacher, 2019). The resultant affective rewards may then be associated with the subject of interaction, leading to a social bonding effect.

Critical to this theoretical framework is that the synchrony-bonding effect may be based upon low-level perceptual processing. Elements of the perceptual experience of synchrony may be common to any synchronised action (Vuoskoski & Reynolds, 2019), as processing fluency is a general aesthetic principle not bound to any sensory modality. As seen in the review of processing fluency literature, a range of judgements may be made about a stimulus based on how easy it is to process, including social judgements (Lick & Johnson, 2015). Stimuli that are easy to recognise and understand are also likely to be familiar, thus fluency may be a useful cue to familiarity. Conversely, disfluency may indicate that a stimulus is unfamiliar, and potentially less safe (Reber & Unkelbach, 2010). Synchrony would appear to fulfil many of the criteria of a fluent stimulus, containing repetitive structures that build familiarity during an extended exposure time, and the effects of synchrony bear a strong resemblance to the effects of processing fluency.

This novel framework remains to be tested. A future research programme should initially be devised to establish whether synchrony truly is easier to perceive and participate in than non-synchrony. The challenges this poses are not insignificant, as synchrony requires continuous action, which may impede efforts to measure reaction times to stimuli. However, if synchrony is easier to process than non-synchrony, then further research could examine whether the social bonding effects of synchrony may in fact be attributable to the experience of fluency, rather than synchrony *per se*. If this is the case, then further questions could be asked about the optimum level of complexity in any interpersonal interaction to achieve such social bonding outcomes, or to build more sophisticated multi-agent models based upon the active inference framework.

Furthermore, the present account may also be reconcilable with pre-existing theories attempting to explain the synchrony-bonding effect at different levels, when considering how social context shapes the experience of synchrony. Shared knowledge, shared goals, and other social factors create the foundation upon which the core processes utilised in synchrony rest (Keller et al., 2014), and a range of contextual cues may inform the predictions one makes about an interaction partner (Veissière et al., 2019), which could influence synchrony. Processing fluency is likely also operating subconsciously (Reber et al., 2004), and the conscious interpretation of its effects may be subject to context. As such, the present account does not seek to supplant previous theories, but to complement them with a perceptual processing account of the experience of synchrony.

Ultimately, this view aims to reconceptualise the synchrony-bonding effect. Rather than relying upon higher-order cognition, the pleasure, sense of social affiliation, increased trust, and bonding that result from synchrony could be – in part – the product of ease of processing. As with processing fluency effects for any stimulus, the experience of fluency arising from synchrony may be used as a cue for familiarity. Rather than creating social bonds *per se*, the

perceiver is deceived into assuming social bonds exist because coordination was easy. The fluency experienced in perception is likely below the level of consciousness and may be subjectively interpreted as a feeling of increased bonding. The function of engaging in synchrony for an individual may then be reconceived as an attempt to make oneself more predictable, to ease processing load for another by creating a seemingly familiar social environment: a manipulation of a simple perceptual mechanism.

1.6 Acknowledgements

The authors wish to thank Susana Carvalho, Martha Newson, Eric Clarke and Bahar Tunçgenç for their helpful comments on early drafts of this manuscript.

2 Processing fluency for visual synchrony perception

Joshua S. Bamford[1], Bronwyn Tarr[1,2], Emma Cohen[1]

[1]University of Oxford

[2]The London Interdisciplinary School

2.1 Introduction

Social behaviour regularly features moments of synchrony that function to display or reinforce social affiliation. Behavioural synchrony is characterised by a strict temporal locking of actions, though the exact shape and kind of the actions may or may not be matched (Bernieri & Rosenthal, 1991). This is exemplified in activities such as music, dance, rowing and marching (McNeill, 1997). Synchrony may emerge spontaneously, as rhythmic sounds often induce involuntary bodily entrainment in adults (Zelechowska et al., 2020b). This capacity for spontaneous sensorimotor synchronisation appears to be limited to a small and distantly related number of animal species, among which humans display particular ability (Fitch, 2013). Despite synchrony's ubiquity, experimental research into has only recently begun to uncover social effects and underlying causal pathways.

Recent studies of the effects of behavioural synchrony on social affiliation and bonding demonstrate that people feel a sense of social closeness and are more likely to exhibit pro-social behaviour towards someone with whom they have previously synchronised – termed the *synchrony-bonding effect* (Hove & Risen, 2009; Mogan et al., 2017; Rennung & Göritz, 2016; Savage et al., 2021; Vicaria & Dickens, 2016). Perceiving others synchronising can also elicit a response: synchronised groups are perceived as more formidable and bonded (Lee et al., 2020), and watching synchronised performances leads to a perception of beauty, aesthetic appreciation and greater enjoyment than non-synchrony (McEllin et al., 2020; Vicary et al., 2017). Participation in, and perception of, synchrony have these social effects, though the mechanisms are poorly understood.

Various causal mechanisms behind synchrony-bonding, and other social effects of synchrony, have been proposed. Some theories rely upon higher-order cognitive processes, such as shared goals, shared intentions, or a perception of self-other similarity (Demos et al., 2012; Reddish et al., 2013; Valdesolo & DeSteno, 2011). Other accounts have suggested that synchrony biases attention to select synchronous stimuli (Bamford et al., 2016; Cirelli et al., 2014; Woolhouse et al., 2016). However, it is not explained why our attention should be drawn to synchrony, nor how synchrony is involved at other stages of stimulus perception.

The causal mechanisms underlying the effects of both participating in and viewing synchrony require further explanation. It is possible that these effects arise through low-level perceptual mechanisms, as synchronised stimuli may be easier to process at the selection and organisation stages of perception, leading to a processing fluency effect (see Chapter 1). Thus, the present study aims to investigate how general perceptual principles contribute to the synchrony-bonding effect.

2.1.1 Processing Fluency

Processing fluency is a general explanatory principle that can be applied to a range of preference judgements. It is not domain-specific and suggests that the ease with which a stimulus is processed provides an intrinsic reward (Winkielman et al., 2003). By rewarding action that seeks out easy experiences, processing fluency effects provide motivation that can function to optimise energy costs of cognition (Leder, 2013). Stimuli that are easy to process evoke positive feelings, which in turn provide information about the stimuli (Schwarz, 2012). Fluency may be a reliable cue for familiarity, without having to rely upon memory processes (Winkielman et al., 2003); one of the most reliable means of generating processing fluency is to have had prior exposure to a stimulus, and so a fluently processed stimulus is likely to be familiar. However, fluency may be experienced because of the intrinsic qualities of the stimulus, not just because of the observer's prior experience (Reber et al., 2004): in this case,

the fluency heuristic may lead to an illusion of familiarity, despite the observer having no prior experience of the stimulus. Thus, one may experience a positive reward from a stimulus that is easy to process – either by the structural simplicity of the stimulus, or through prior exposure – which then implies familiarity.

Processing fluency is a general principle that has been successfully applied to a wide range of stimuli. It has been used to explain preferences for art, visual patterns, music and font styles (Reber et al., 2004). Processing fluency has even been applied to some social judgements, as people are more likely to be prejudiced against others who are harder to understand (Lick & Johnson, 2015). Despite its generality, processing fluency has yet to be tested for synchronised stimuli. In a purely theoretical exploration, Wheatley and Sievers (2016) suggested that interpersonal synchrony may be more efficient to process by consolidating streams of information about the self and other, leading to a state of fluency. Hove and Risen (2009) also raised the possibility that synchrony may be easier to process than non-synchrony. The suggestion was offered in a brief footnote and has not since been theoretically developed or tested.

Despite this lack of empirical testing, many observations associated with the synchrony-bonding effect and the perception of synchrony bear similarity to the effects of processing fluency. Moving in synchrony with another increases their likability (Hove & Risen, 2009). Meanwhile, greater synchrony in performance increases viewer enjoyment and aesthetic ratings (McEllin et al., 2020; Vicary et al., 2017). Similar effects on liking and aesthetic perception are observed in typical processing fluency studies (Reber et al., 2004). The structure of synchronised stimuli may fit within the existing criteria for fluent processing. Synchrony is usually sustained over an extended period of time (Bispham, 2018), and repeated exposure to a stimulus builds familiarity and fluency for that stimulus (Reber et al., 1998; Winkielman & Cacioppo, 2001). Fluently processed stimuli often have simple features and contain a degree

of redundancy, such as in symmetrical patterns (Makin et al., 2012). This could be analogous to the repetitive movement patterns usually observed in synchronised action, which often exhibit a simple rhythmic structure (Polak et al., 2018). Synchronised stimuli share traits with other stimuli that are easy to process, and the resultant processing fluency potentially produces positive affective and social consequences for the perceiver(s).

Synchronous action entails the grouping of visual, auditory or other sensory objects (Hukin & Darwin, 1995; Rideaux et al., 2016). For example, when two people tap or drum together in synchrony, the listener may perceive the two objects and combine them together to form a larger whole. Wheatley and Sievers (2016) described a phenomenon of interpersonal streaming which may occur during synchrony. As such, Gestalt grouping principles can inform the perceptual processes involved.

Gestalt grouping, first described by Wertheimer (1938), outlines a set of principles through which objects are grouped together. Cues such as proximity and similarity result in many objects being seen as part of a larger whole. The principles apply to both visual and auditory perception, although the process of grouping sounds into ‘streams’ is usually referred to as Auditory Scene Analysis (Bregman & Campbell, 1971). Crucially, synchrony of sound or movement, or of the onset of an event, is a strong cue for grouping in both auditory and visual domains (Hukin & Darwin, 1995; Rideaux et al., 2016). When multiple stimuli act in synchrony with each other, they could be perceived as one object due to gestalt grouping principles or auditory streaming depending on the modality (Hukin & Darwin, 1995; Pressnitzer et al., 2011; Rideaux et al., 2016; Wertheimer, 1938). However, ambiguity in grouping cues requires increased attentional effort to resolve (Dolležal et al., 2014). In the acoustic domain, additional perceptual objects provide a distraction, demanding more attentional control (Hughes et al., 2013; Zeamer & Fox Tree, 2013). These grouping principles

then provide a strong candidate for explaining why synchrony could be easier to process than non-synchrony.

2.1.2 The Present Study

In order to understand the causal factors underlying the synchrony-bonding effect, the present study examines the perceptual experience of synchrony. It builds upon the suggestion by Wheatley and Sievers (2016) that synchrony may be enjoyed because it is easier to process than non-synchrony. Specifically, this study tests whether synchrony is easier to process than non-synchrony.

Ease of processing is regularly measured using reaction times (Reber et al., 2004). However, prior research has found that participants may choose to sacrifice speed for accuracy (Dutilh et al., 2011). This speed-accuracy trade-off may be accounted for by instructing participants to focus on either speed or accuracy, or by combining speed and accuracy into an overall performance score. The present study used a performance score, combining reaction time and error rate, to infer the processing load of the tasks, while also instructing participants not to sacrifice accuracy in their responses.

Reaction times have rarely been studied in the context of synchronised action because synchrony requires precise timing by its nature. Reaction times can be influenced by spontaneous entrainment to a periodic stimulus, as people will tend to respond in time with the stimulus (Killingly et al., 2021). Therefore, the present study uses non-periodic stimuli, involving continuous movement without any rhythmic structure.

Most synchrony-bonding studies have used periodic stimuli. In periodic movement, velocity changes in a regular, oscillating pattern. When two oscillators are in synchrony, their velocities will always be changing, but will have a consistent relationship between them. By using non-periodic stimuli with a consistent velocity, the present study avoids spontaneous

entrainment to the stimuli. As a consequence, it focuses on the perception of brief snapshots of coordinated movement between multiple stimuli.

The stimuli used in the present study were based upon the animated Flanker paradigm developed by Lange-Malecki and Treue (2012), but with Gabor patches similar to those used by De Valois and De Valois (1991). Gabor patches are generated through the convolution of sinusoidal and Gaussian functions (Figure 2.1). They can be animated so that the bars appear to drift and are often used to study motion perception (e.g. Amano et al., 2009; De Valois & De Valois, 1991; Rideaux et al., 2016). Previous studies have suggested that animated Gabor patches moving in synchrony are perceptually grouped together (Rideaux et al., 2016). Animated Gabor patches are thus ideal stimuli for studying processing fluency that may arise through perceptual grouping.

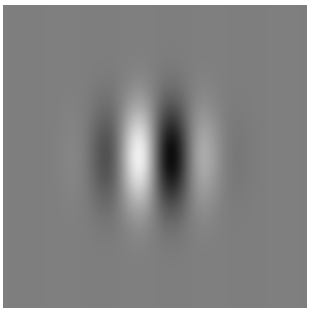


Figure 2.1 A single Gabor patch.

Developed by Eriksen and Eriksen (1974), the Flanker task is a method of studying attention control. Typical flanker tasks involve a central ‘target’ stimulus, surrounded by ‘flanker’ stimuli in a horizontal line either side. The stimuli are usually static images, such as a letter or an arrow, that correspond to a keyboard response. In congruent conditions, the ‘target’ and ‘flankers’ are identical, while in incongruent conditions they are not. Participants complete many trials, with the average reaction time and accuracy of responses to the stimuli

taken as measures of the effect of visual distraction between congruent and incongruent conditions (Eriksen & Eriksen, 1974). Prior research has demonstrated that reaction times can be reliably measured with Flanker tasks on an online platform (Crump et al., 2013; de Leeuw, 2015; Pinet et al., 2017).

Although a well-established test for the perception of static images, there are few examples of animated Flanker tasks. One study, with a relatively small sample of 21 students, directly compared the effects of moving and static flankers (Lange-Malecki & Treue, 2012). In all conditions, participants had to report the direction indicated by the target, but in the static condition the stimuli were all triangles pointing either left or right, while in the moving condition the stimuli were random dots moving within a defined circle. The authors found similar effects of congruency in both of these conditions (Lange-Malecki & Treue, 2012). However, the aim of their study was not to study synchrony *per se*, and their findings were based on a very small sample. Lange-Malecki and Treue (2012) did demonstrate that the flanker effect exists for moving as well as static stimuli, which the present study investigates further with the specific aim of testing the processing load of synchrony.

Across two experiments, we tested the hypothesis that synchronised stimuli are easier to process, as measured by performance on a Flanker task, than non-synchronised stimuli. Study 1a used only one manipulation, having stimuli that were either congruent or incongruent in direction. Study 1b built upon Study 1a through the addition of a speed manipulation, expecting highest performance when the target was moving at the same rate and direction as the flanker stimuli. Both experiments were conducted online using the novel Gabor-Flanker task. This is a variation on the animated Flanker task developed by Lange-Malecki & Treue (2012), although featuring animated Gabor patches such as those used by De Valois and De Valois (1991). Better performance in the congruent trials would be taken as support for our hypothesis of decreased processing load due to the target and flanker stimuli being perceptually

grouped together. Lower processing load when perceiving synchronised stimuli than when perceiving non-synchronised stimuli would suggest that some of the positive effects of synchrony could, at least in part, be explained by processing fluency.

2.2 Study 1a

In a within-subjects design, Study 1a tested the effect of synchrony on processing load, as measured by reaction times and a composite performance score. It utilised novel stimuli by constructing a Flanker task using Gabor patches, in which the movement of the target patch may be either congruent or incongruent with the movement of the flankers by moving either in the same or opposite direction (Lange-Malecki & Treue, 2012). In line with our hypothesis that synchrony is easier to process than non-synchrony, we predicted that participants would make faster and more accurate responses when the flanker stimuli appeared to be moving congruously (i.e. in the same direction) with the target/central stimulus, compared with when they appeared to move incongruously (i.e. in different directions). The study was pre-registered.[2]

2.2.1 Participants

This study was conducted online, with participants recruited using snowball sampling methods through social media (Facebook, Twitter, Reddit). A sample of 73 participants (45 female, 27 male, 1 other) completed the study, exceeding the target sample size of 71 (based upon a power analysis with d = .3, α =. 05, 1-β = .8). Most participants were living in Western countries, with the majority coming from the United Kingdom (34.2%), the rest of Europe (30.1%), Australia (23.3%), and North America (8.2%), and had a mean age of 31.01 (SD = 10.3, range = 20:62). Participants were required to have normal or corrected vision, and to have access to a device with a keyboard. All participants reported fulfilling these criteria, and gave

[2] http://aspredicted.org/blind.php?x=cj7bd7

informed consent to participate in this study, in accordance with the ethics requirements of the university's ethics board. See Appendix 6.1.1 for the information sheet and consent form.

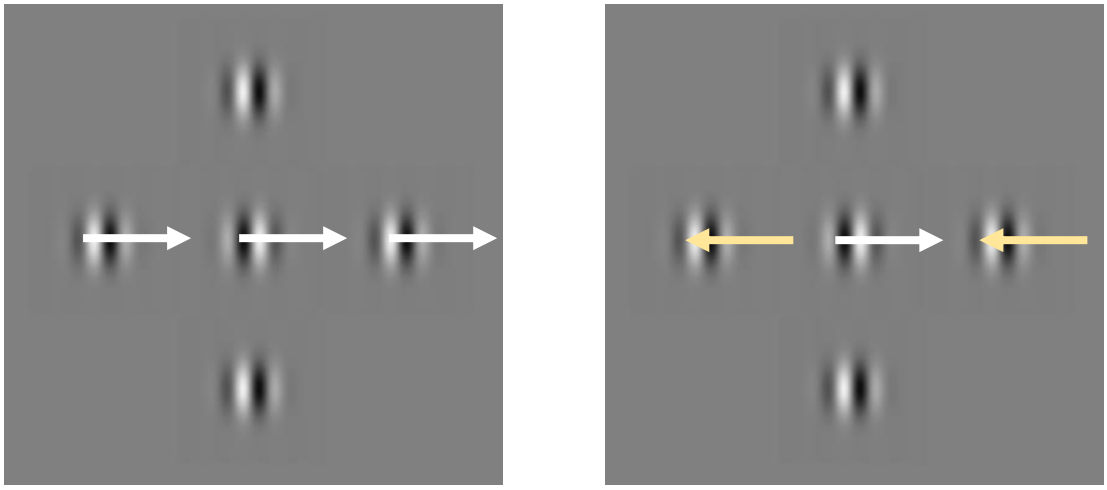
Any incomplete responses were excluded. Mean time to complete the perceptual component of the study (excluding the demographic questionnaire) was 452.82 seconds (*SD* = 112.62). One participant was excluded from this calculation for having an extraordinarily long completion time (12296.4 seconds), although most of this time was taken on the first screen, and there was no unusual behaviour otherwise on the task. It is suspected that this participant left the page open before beginning the task itself, and thus their data was retained for analysis.

2.2.2 Materials

The online study consisted of two parts: a short demographic questionnaire – see Appendix 6.1.2 – and the Gabor-Flanker task. The animated Gabor stimuli in this study were generated in MATLAB using a custom function.[3] Three Gabor patches were generated in this way: static, left-movement, right-movement. Each individual Gabor patch was 150 pixels in diameter, with a period length of 1cm, and the animated patches moved at a rate of 1cm/s. During development of the task, pilot participants had reported feeling disoriented when all the stimuli were moving, so static patches were added above and below the target to act as a visual reference point. The Gabor-Flanker task presented five Gabor patches (one target, two flankers, two static references) arranged in a diamond (Figure 2.2). There were two conditions in a within-subjects design: a congruent (synchrony) or incongruent (non-synchrony) condition. In the congruent condition, the target Gabor moved in the same direction as those either side of it, and in the incongruent condition it moved in the opposite direction to the flankers. The Gabor patches immediately above and below the target were always static. The task itself was built using jsPsych (de Leeuw, 2015), and the script is available in the supplementary online

[3] https://github.com/bamford-js/GaborGen

materials.[4] Because the task was completed by participants online, on their own computers, specific details of resolution and screen size could not be controlled.



a) Congruent b) Incongruent

Figure 2.2 Arrangement of stimuli in the Gabor-Flanker task. The central Gabor was always the target stimulus, and the patches immediately above and below it were always static, while those to its left and right were animated. The flanker stimuli either moved in the same direction as the target (Congruent) or in the opposite direction (Incongruent). Arrows for illustration purposes only.

2.2.3 Procedure

Participants completed the demographic questionnaire in Qualtrics, before being presented with the Gabor-Flanker task. The on-screen instructions explained that this study aimed to investigate motion perception. Specific details of the hypotheses were not divulged. Participants were told to categorise the direction of the target (either left or right) and were instructed to respond to each trial as quickly as possible, without sacrificing accuracy. Previous studies have found that participants adopt two different modes of thinking in reaction time experiments: stimulus-controlled behaviour, and guessing (Dutilh et al., 2011). By emphasising the need for accuracy, guessing can be reduced.

[4] https://github.com/Social-Body-Lab/fluency-through-synchrony/tree/main/Study1/Study1_Experiment

The Gabor-Flanker task was initially presented in 8 practice trials, followed by three blocks of 60 trials. Participants were invited to take a short break between blocks, to prevent fatigue. Each block contained 30 trials from each condition (congruent and incongruent), presented in randomised order. Participants had a maximum of 2000 ms in which to respond to each trial, or that trial would be marked as incorrect. Responses were given by pressing either the 'F' or 'J' key on their keyboard, for Left and Right respectively, to indicate the direction of the motion in the target Gabor patch. Following every trial, there was a short pause of between 500 and 2000 ms in duration. The random pause duration was to prevent the task from having any intrinsic rhythm. The script may be found in Supplementary Materials 3.

2.2.4 Analysis

Raw data were exported from Qualtrics. Reaction time (RT) – the time from initial presentation of a stimulus until the participant's button press – was recorded for each trial within the online JavaScript application. For each block of 30 trials, a mean reaction time for correct responses was calculated, as well as an accuracy rate (percentage of correct responses in a block). Accuracy (%) was divided by mean RT (ms) on correct trials to give a composite performance score (hereafter 'performance'). This was calculated for each participant, in each block, for every condition, and then scaled by 100 for readability. This is an adaptation from Lange-Malecki and Treue (2012), who used an inverse efficiency measure to account for different accuracy rates between conditions. Although participants were instructed not to sacrifice accuracy, it is possible some may have done so to optimise RT. Therefore, RTs alone may not capture the true difficulty of the task. As a combined measure, performance accounts for any potential speed-accuracy trade-off. A Pearson's correlation was used to confirm a negative relationship between accuracy and RT.

To test the hypothesis that the non-synchronised (incongruent) trials are more difficult to process than synchronised (congruent) trials, a repeated measures ANOVA was used to test

the effect of flanker condition (stimulus) and presentation order (block) on performance. Presentation order was included to investigate any practice effects.

Data were processed in Python, statistical tests were conducted using Pingouin (Vallat, 2018), and figures were produced with Seaborn (Waskom, 2021). The data and script for analysis may be found in the supplementary online materials.[5]

2.2.5 Results

Descriptive statistics for all measures on the stimulus conditions are reported in Table *2.1*. Participants who scored more than 2.5 standard deviations from the mean on performance were removed as statistical outliers. The level of 2.5 standard deviations has been used in previous processing fluency studies as the basis for exclusion (Reber & Schwarz, 1999). Five participants were removed in this way. These same five participants were observed to have an accuracy rate below chance level (50%) in at least one condition of at least one block of trials. In every case, this reduced accuracy rate was in the incongruent condition (see Figure *2.3*).

A medium, positive, significant correlation was observed between mean reaction time and accuracy, $r = 0.277$, $p < 0.001$, indicating a trade-off between reaction time and accuracy (see Figure *2.4*), and justifying the use of a composite performance score to account for this.

Table 2.1 *Descriptive statistics for reaction time, accuracy, and composite performance across the two stimuli conditions, with outliers excluded.*

	Reaction Time (ms)		Accuracy (%)		Performance (%/ms)	
	M	*SD*	*M*	*SD*	*M*	*SD*
Congruent	573.6	100.2	97.76	4.142	17.46	2.605
Incongruent	578.0	105.2	97.70	3.699	17.32	2.460

[5] https://github.com/Social-Body-Lab/fluency-through-synchrony/tree/main/Study1/Study1_Analysis

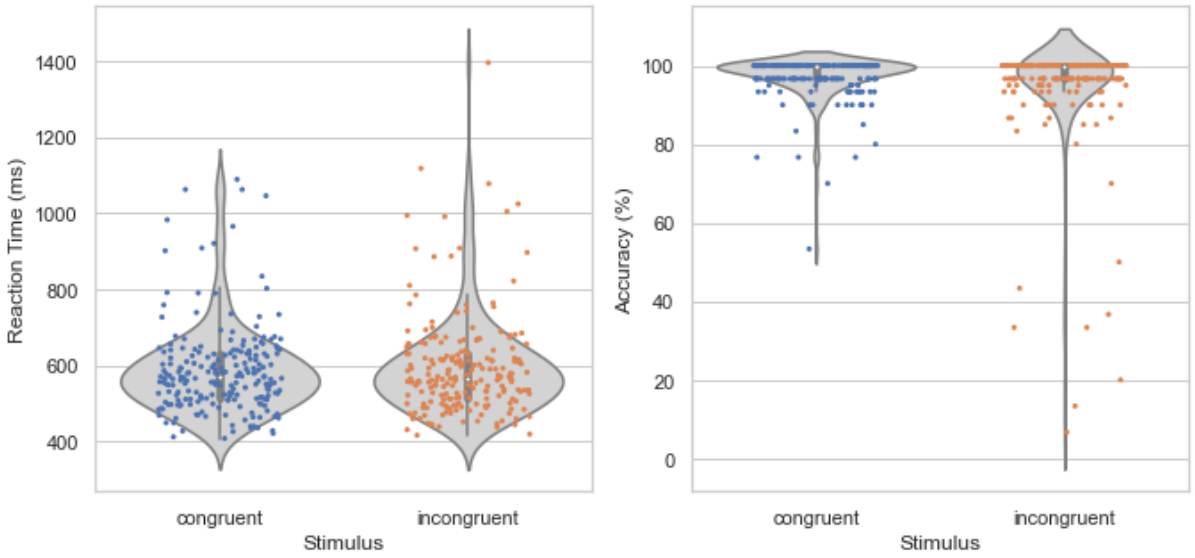


Figure 2.3 Reaction time (s) on correct trials, and accuracy (%) per block, over Flanker condition (stimulus type). Includes outliers, showing those with low accuracy on incongruent trials.

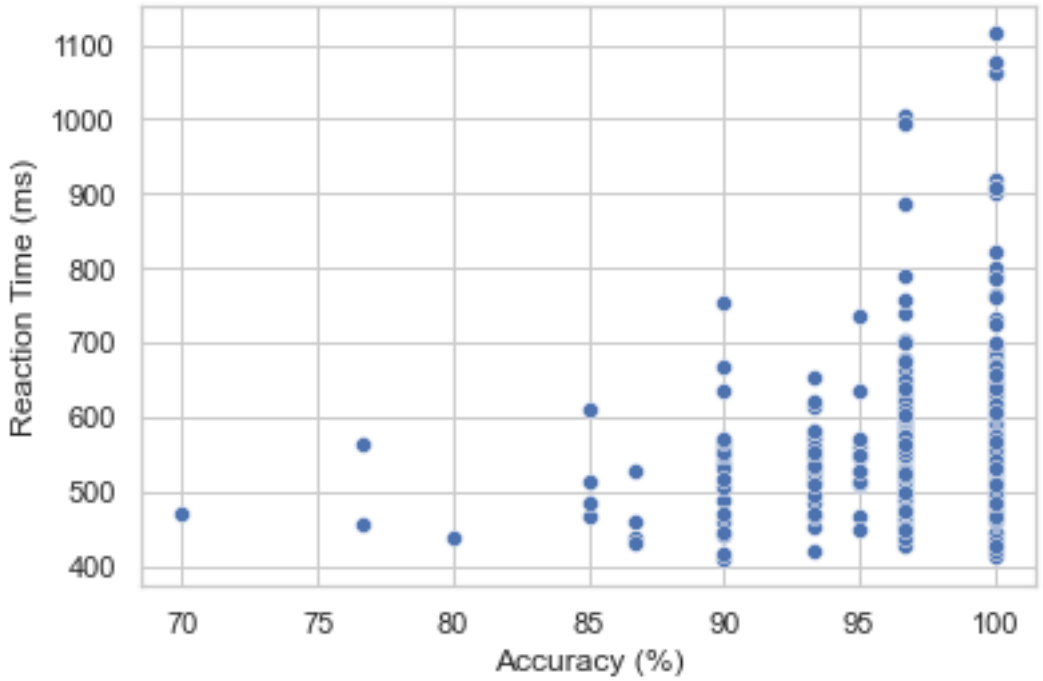


Figure 2.4 Mean reaction time (ms) on correct trials per block, plotted against accuracy (%) per block.

Using the composite performance measure (accuracy/reaction time), the assumptions of equal variances were met for both the stimulus condition, $W = 1.04$, $p = .309$, and block condition, $W = 2.270$, $p = 0.105$. However, the assumption of sphericity was violated for the block condition, $X^2(2)= 23.76$, $p < .001$, so a Greenhouse-Geisser correction was applied. The repeated measures ANOVA found a significant effect of block on performance, but no effect of stimulus condition (see Table 2.2 and Figure *2.5*). Post-hoc, Bonferroni corrected t-tests revealed significant differences between Block 1 and Block 2, $t(66) = -4.842$, $p < 0.001$, and between Block 1 and Block 3, $t(66) = -4.138$, $p < 0.001$, suggesting a significant practice effect. There was no significant effect of stimulus condition on performance.

Table 2.2 *Repeated-measures ANOVA for the effect of stimulus condition and block number on performance.*

Effect	*df*	*F*	*p*	η_p^2	ε
Stimulus	1, 66	1.469	.230	0.022	1
Block	2, 132	16.30	<.001	0.198	0.766
Stimulus * Block	2, 132	0.529	.578	0.008	0.930

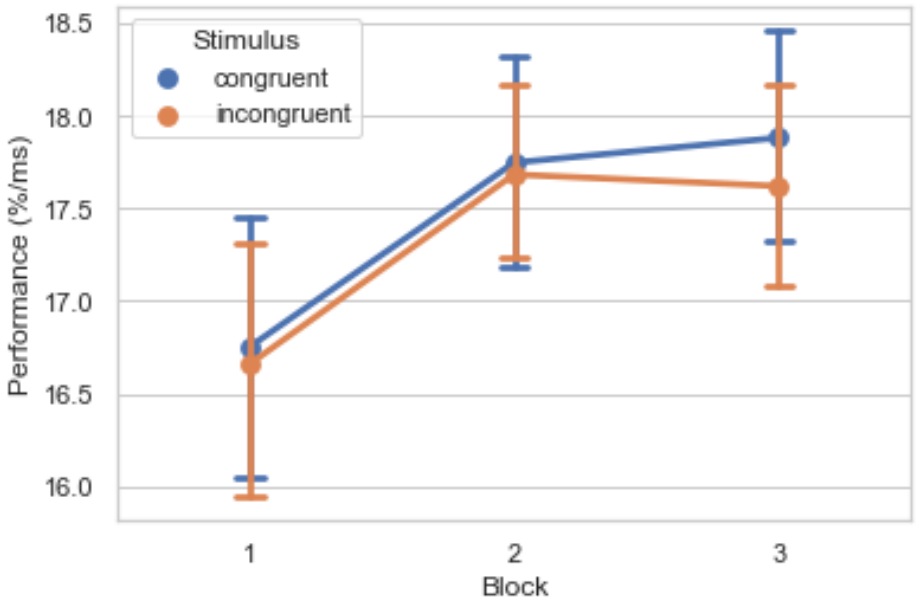


Figure 2.5 Plot of each participant's composite performance score (accuracy/reaction time) in both the congruent and incongruent conditions, against presentation order (block number). All data are presented, including subjects that scored below chance level accuracy. Error bars show 95% confidence intervals.

2.2.6 Discussion

Study 1a tested the effect of synchrony on perceptual fluency, operationalised as a performance measure combining both accuracy per block and reaction time (RT) for correct trials. In this case, synchrony within a trial was manipulated through congruency of direction between the target and flanker stimuli. Trials were divided into three blocks, for the purpose of calculating accuracy rates and to investigate order effects. Significant order effects were observed, with participants showing increased performance on later blocks. This indicates a practice effect, which is consistent with previous literature on attention control tasks such as the Flanker (Chen et al., 2013). However, the observed effects of the stimuli were more mixed.

No effect was found of stimulus condition on performance. It should be noted, however, that there were five outliers excluded for being greater than 2.5 standard deviations from the mean. These participants also scored below 50% accuracy (chance level) on at least one

condition within at least one of the blocks. In every case this was only for the incongruent condition.

Although, on average, there was no effect of condition, the performance of a subset of participants – those five outliers – did appear to be impacted. Given that the study was conducted online, and the participants' behaviour during the task could not be observed, it is possible that they had either misunderstood the task or were distracted. However, the presentation of stimulus type was randomised for each trial, and the increased rate of errors in the incongruent condition seems systematic for some participants. This could be because they found these trials more difficult to perceive, failing to correctly identify the direction of motion. All participants should have had normal or corrected vision, but it is possible that some participants had a vision or attention impairment that they either did not disclose or of which they were not aware. Alternatively, some participants potentially experienced something akin to the wagon wheel illusion, in which a stimulus appears to move backwards because of a mismatch between the rate of movement and the frame rate of the video (Purves et al., 1996), although the frame rate of modern computer monitors (usually at least 50Hz) should preclude this. Lange-Malecki and Treue (2012) also observed a reduced accuracy rate for incongruent moving flankers, although they did not report whether this disproportionately affected some participants more than others.

Typically, a congruency effect is observed in Flanker tasks, with reaction times being faster when the target stimulus is congruent with the flankers (Chen et al., 2013; Eriksen & Eriksen, 1974), and this has previously been observed in animated Flanker tasks similar to those in the present study (Lange-Malecki & Treue, 2012). Study 1a failed to replicate the results of Lange-Malecki and Treue (2012), though differences in design must be acknowledged. In both studies, the incongruent conditions were incongruent only in direction. However, Lange-Malecki and Treue (2012) used stimuli with moving dots, rather than the

animated Gabor patches presented here. The dots used by Lange-Malecki and Treue (2012) were more visually complex, appearing as more objects on the screen, and thus less susceptible to being grouped together. Conversely, two Gabor patches are likely to be perceived as one object (Rideaux et al., 2016), even with opposing directions of motion. Because of this general difference in visual complexity, incongruency in Gabor direction may not have been difficult enough to register a difference in reaction time, when compared to incongruency in dot movement.

The null finding here, with stimuli that exhibit a consistent velocity, may also be understood in terms of periodic motion. In natural perception of synchrony, objects may be moving in different directions and/or at different speeds at any given moment. Natural oscillators exhibit periodic movement in which direction changes at regular intervals, while the stimuli presented here maintain a consistent direction. Despite the Gabor patches themselves being sinusoidal, their velocity remains consistent, and the relationship between the target and its flankers is constant within a trial.

In synchrony studies, antiphase coordination is usually still considered to be synchronised (Repp, 2005). This occurs when two agents have a stable temporal relationship, except that one agent reaches the peak of its motion as the other is at its trough. The derivative (velocity) of these two waves would always be moving in opposite directions, at the same speed: just as the incongruent condition presented here. Therefore, incongruency in this experiment could be considered to be a derivative of antiphase synchrony, and thus not a true representation of the perceptual experience during a moment of non-synchrony.

When two oscillators move in synchrony, they will always be matched in velocity at any timepoint. Study 1a had one incongruent condition, in which direction was opposing but speed remained matched. In non-synchronised movement between two oscillators, speed

usually also differs at any given time between them. To investigate the effect of both speed and direction, Study 1b includes a speed manipulation as an additional form of incongruency within the Flanker paradigm.

2.3 Study 1b

Study 1b aimed to replicate and extend the results of Study 1a by examining the effect of a speed manipulation. In Study 1a, the target was either moving in the same, or opposite, direction to the flankers. Study 1b also varied the rate of movement between the target and flankers in a 2x2 within-subjects design, with each participant completing four conditions: speed congruent – direction congruent (SC-DC), speed incongruent – direction congruent (SI-DC), speed congruent – direction incongruent (SC-DI), and speed incongruent – direction incongruent (SI-DI).

Interpersonal synchrony can be better tested in this reduced form by a speed manipulation than by the direction-based manipulation of Study 1a. It was hypothesised that speed-incongruency would have a negative effect on performance. Based upon the results of Study 1a, no effect of direction-incongruency was expected. There were no specific hypotheses about interactions between the two factors. On one hand, there could be an additive effect of increased processing difficulty when both factors are incongruent. On the other hand, the target may become so different from the flankers, based on Gestalt grouping principles, that it becomes easier to perceive separately. This experiment was pre-registered.[6]

2.3.1 Participants

Participants were recruited online through the same methods as Study 1a, however with an entirely new sample (no participants from Study 1a took part in Study 1b). The minimum sample size was 62, for a repeated-measures ANOVA of small-medium effect size ($\alpha = .05$, 1-

[6] https://aspredicted.org/blind.php?x=wm8g3c

β = .8). 83 participants completed the task (49 female, 30 male, 4 other; mean age of 31.73). The majority of participants were from the UK (45.8%), the rest of Europe (32.3%), Australia (10.8%), North America (8.4%), and Asia (2.4%). Study 1b had the same selection criteria as Study 1a. All participants reported having normal or corrected vision and gave informed consent according to the university's ethics requirements. See Appendix 6.1.1 for participant information and consent forms. This study was approved by the Departmental Research Ethics Committee as an extension of Study 1a (Ref: SAME_C1A_19_002).

Mean completion time for the perceptual component of the study (excluding the demographic questionnaire) was 638.70 seconds (*SD* = 174.37). Two participants were excluded from this calculation for having unusually long times (2279.14s and 2764.98s) although the excessive time was spent on the first screen, and otherwise they completed the task as expected, so their data were retained for analysis. Incomplete responses were excluded.

2.3.2 Materials

Study 1b consisted of the same demographic questionnaire (Appendix 6.1.2), and a similar Gabor-Flanker task to that used in Study 1a. Four Gabor patches were generated using the same MATLAB script as Study 1a, with both left and right animated patches at both slow and fast speeds for each condition. The fast Gabor patches moved at 1cm/s, the same rate as those in Study 1a, while the slow patches moved at 0.5cm/s. The precise parameters to generate these Gabor patches may be found in the code within the supplementary online materials.[7] Each patch was 150 pixels in diameter. These were then arranged into the same diamond-shaped configuration as Study 1a (see Figure 2.2), although with the added speed manipulation. As with Study 1a, the Gabors on the vertical axis were held static to use as a visual anchor. Four conditions (as listed above) were constructed based on the relationship between the target

[7] https://github.com/Social-Body-Lab/fluency-through-synchrony/blob/main/Study1/GaborGen.m

Gabor in the centre and the two flankers to the left and right. All possible variations of speed and direction existed within these conditions, e.g. target slow and moving leftwards, flankers fast and moving rightwards would be a SI-DI stimulus.

2.3.3 Procedure

Participants took part in the study remotely. They were instructed to use a home computer with access to a keyboard, in a quiet environment, free from distractions. Participants completed the short demographic questionnaire (as in Study 1a) in Qualtrics. This was followed by the perception task, which appeared embedded as a JavaScript App within Qualtrics. The script may be found in the supplementary online materials.[8] Participants were given similar instructions to Study 1a.

There was a practice block of 16 trials, followed by 5 blocks of 48 experimental trials. This resulted in 60 experimental trials in each of the 4 conditions, presented in a randomised order, equally divided between each block. At the start of each trial, a Gabor-Flanker grid was presented and then participants had up to 2000ms to respond to each trial, after which that trial would be marked incorrect. As in Study 1a, there was a gap between trials of between 500ms and 2000ms (duration randomised). A hypothesis probe was included at the end, which was not done for Study 1a.

2.3.4 Analysis

As in Study 1a, we calculated a composite performance score (accuracy divided by mean reaction time for correct trials) for each participant in each condition. Two-way repeated measures ANOVAs were used to analyse the effect of the two flanker conditions (speed and direction) on composite performance. The composite performance measure was used due to the different accuracy rates between conditions, as in Study 1a. To examine possible

[8] https://github.com/Social-Body-Lab/fluency-through-synchrony/tree/main/Study1/Study1_Experiment

confounds, a second 2x2 repeated measures ANOVA tested for the effect of target speed (the speed of the central Gabor patch) against the flanker effect of speed (the difference between the speed of the central Gabor and that of the flanker Gabor patches). Results were processed in Python; statistical tests were conducted using Pingouin (Vallat, 2018), and figures were produced with Seaborn (Waskom, 2021). The data and script for analyses may be found in the supplementary online materials.[9]

2.3.5 Results

Descriptive statistics for all conditions may be found in Table *2.3*. Initial analysis revealed that five participants scored below chance level (50% accuracy) on at least one condition in at least one block. In all cases these were in either the SC-DI or SI-DI condition (see Figure 2.6). In the free-response questions at the end of the study, one of these participants reported that they sometimes perceived the stimuli as 'rotating but not moving' and so did not respond to those trials. This could have been grounds for removing these participants as outliers, however doing so did not change the results significantly, and so we retained them in the final sample for analysis.

[9] https://github.com/Social-Body-Lab/fluency-through-synchrony/tree/main/Study1/Study1_Analysis

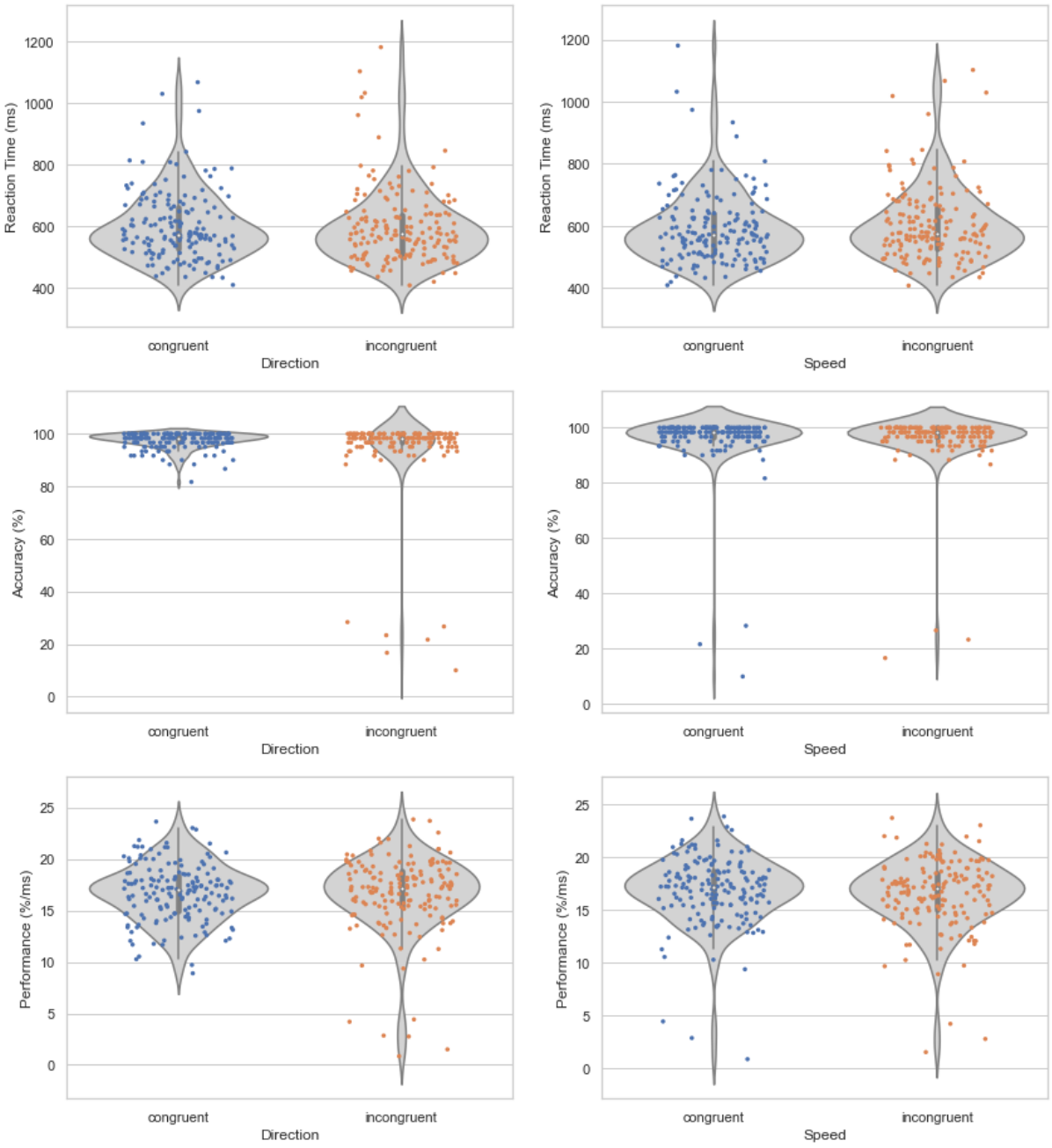


Figure 2.6 Mean reaction time, accuracy and performance per condition. Each point represents a participant per condition, averaged across trials per condition.

As discussed in Study 1a, composite performance scores are used to control for any confounding effects of a speed-accuracy trade-off. Performance scores were calculated in the same manner as Study 1a. A two-way repeated-measures ANOVA was conducted to examine condition-wise effects on performance. The assumption of sphericity was met, as both the speed and direction factors only had two levels (congruent and incongruent). Both factors also

met the assumption of equal variances (direction: $W = 0.161$, $p = .688$; speed: $W = 2.412$, $p = .120$). Significant main effects were found for speed and on the interaction between factors, but not for direction (see Table *2.4* and Figure 2.7). Post-hoc analysis with Bonferroni corrected t-tests revealed significantly better performance in the speed-congruent (SC) trials, compared with speed-incongruent (SI) trials, $t(82) = 3.513$, $p < .001$. Furthermore the effect of speed was more pronounced in the direction-congruent (DC) than the direction-incongruent (DI) trials, such that performance was highest in the DC-SC trials and lowest in the DC-SI trials, $t(829) = 4.567$, $p < 0.001$.

Table 2.3 *Descriptive Statistics for mean reaction time, accuracy and performance across all conditions on both congruency factors and for speed of target.*

		Reaction Time (ms)		Accuracy (%)		Performance (%/ms)	
Condition		*M*	*SD*	*M*	*SD*	*M*	*SD*
Speed	Congruent	599.5	132.8	95.74	12.99	16.68	3.535
	Incongruent	609.9	130.7	95.64	12.47	16.37	3.547
Direction	Congruent	604.9	122.7	97.48	3.020	16.66	2.865
	Incongruent	604.5	140.4	93.90	17.57	16.39	4.109
Target	Faster	577.8	122.3	96.47	11.34	17.36	3.517
	Slower	628.1	138.3	95.20	13.45	15.83	3.578

Table 2.4 *Results of two-way repeated-measures ANOVA of speed and direction factors on performance scores.*

Factor	*df*	*F*	*p*	η_p^2
Speed	1, 82	12.51	<.001	0.132
Direction	1, 82	0.918	.341	0.011
Speed * Direction	1, 82	7.214	.008	0.081

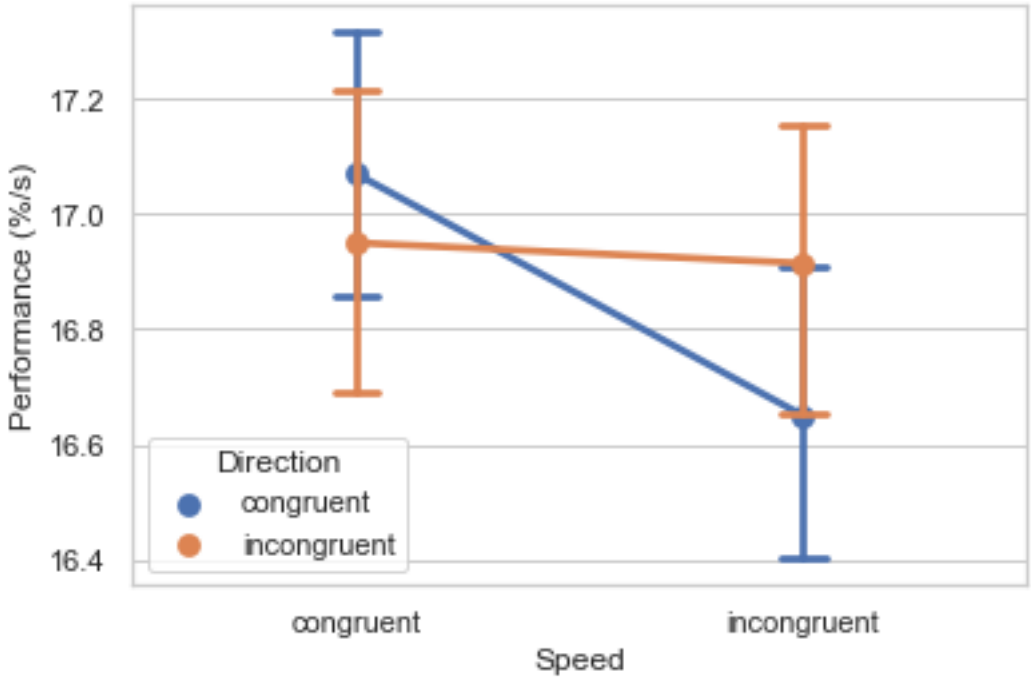


Figure 2.7 Interaction between the speed and direction flanker conditions. Error bars show 95% confidence intervals.

A further repeated-measures ANOVA was conducted between speed congruency and the speed level of the target Gabor. This was an exploratory analysis, to test for a possible confound between these two factors, and had no preregistered hypothesis. The assumptions of sphericity and equality of variances were met (target speed: W = 0.233, p = .629; speed congruency: W = 2.412, p = .120). Significant independent effects were found in both target speed and speed congruency (see Table *2.5*). Post hoc analyses found performance to be significantly higher when the target was fast compared with slow, $t(82) = 18.08$, $p < .001$, and when speed was congruent compared with incongruent, $t(82) = 3.513$, $p < .001$. An interaction effect was also observed within the slow trials, with performance being higher in the slow-SC condition compared with the slow-SI condition, $t(82) = 3.467$, $p = .001$. No such effect was observed within the faster condition, $t(82) = -0.792$, $p = .856$ (see Figure 2.8).

Table 2.5 *Results of two-way repeated-measures ANOVA of speed and direction factors on performance scores.*

Factor	*df*	*F*	*p*	η_p^2
Speed Congruency	1, 82	11.73	<.001	0.125
Target Speed	1, 82	315.0	<.001	0.793
Congruency * Target	1, 82	20.77	<.001	0.202

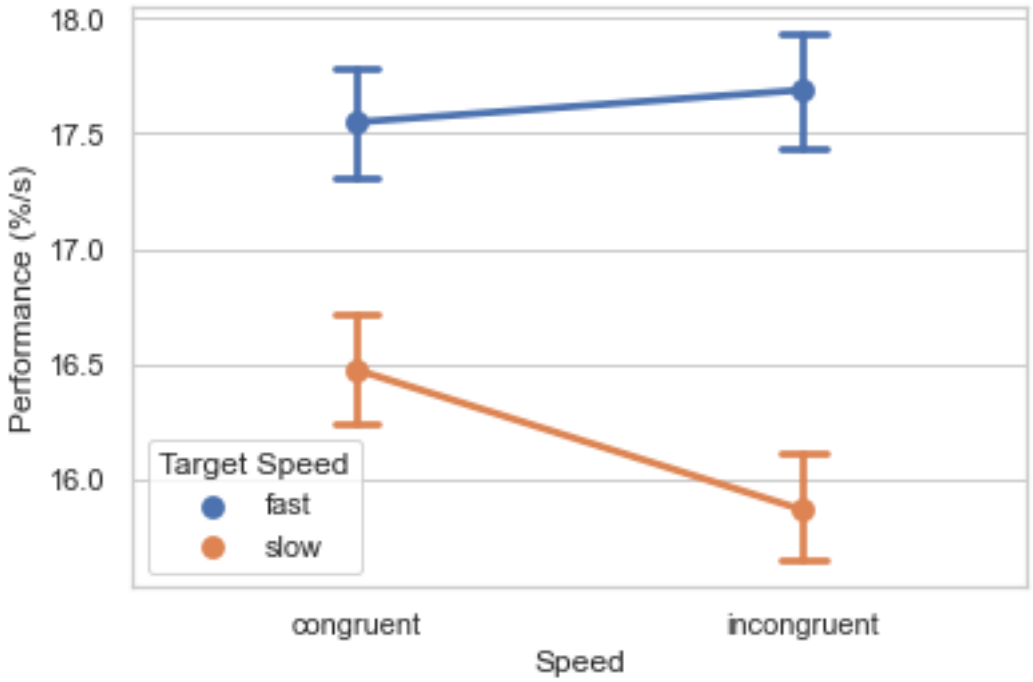


Figure 2.8 Interaction between target speed and speed congruency. Error bars show 95% confidence intervals.

2.3.6 Discussion

Study 1b was designed to replicate and extend the findings of Study 1a by including two flanker conditions: speed and direction. In each trial the target stimulus moved either in the same or different direction and either at the same speed, faster or slower than the flanker stimuli. As with Study 1a, although there was a small number of participants who had severely reduced accuracy in the direction-incongruent conditions, no effect of direction-congruency was detected on overall performance.

Analyses revealed a significant effect of speed on performance. Task performance was lower when the flankers were moving at a different speed to the target. Although only a relatively small increase in reaction time was observed due to incongruent speed, when combined with error rate in the overall performance score this became a moderately large effect. This suggests that differences in speed between two stimuli may influence perception of those stimuli, making them more difficult to process, consistent with our account that synchrony is associated with processing fluency. Further investigation found that this reduced performance for a speed-incongruent flanker trial occurred primarily in the trials that were also direction-congruent. Temporal synchrony, as occurs in the speed-congruent condition, can be used as a grouping cue (Rideaux et al., 2016). When the flankers are moving in the opposite direction to the target this may make the target more salient, creating a stronger figure-ground contrast (Wertheimer, 1938). The greater contrast between the target stimulus and the background may then make the target easier to perceive (Kornmeier & Bach, 2012). If the decrease in performance on speed-incongruent trials is due to ambiguity between the target and flankers, then adding an additional direction-incongruence may serve to ameliorate that ambiguity, making the target easier to perceive.

To check for any possible confounds, the speed of the target was examined. There was an independent effect of the target's speed relative to the flankers, with participants performing worse on slower (vs faster) trials. However, an interaction effect was also observed, showing a greater performance decrease when the target was slower on incongruent trials (when the flankers were moving faster). Nothdurft (2002) found that, when stimuli are moving at different speeds, faster motion is more salient – more readily capturing visual attention. In the context of our experiment, faster stimuli may be more salient. When the target is moving at a faster speed, this improves performance, however if the flankers are moving faster than the target,

they provide a greater distraction, and performance is impacted. Conversely, there seemed to be no effect of the flankers when they moved slower than the target.

From this study, it appears that incongruent movement between non-periodic stimuli may be more difficult to perceive than congruent movement, but particularly when the target was slower than distractor stimuli. The faster stimulus may be more salient (Nothdurft, 2002), thus extra effort is required to direct attention towards a slower target. As in Study 1a, there was no effect of direction in Study 1b. The different effects of speed and direction in Study 1b may be pertinent when considering the implications of these results for the perception of periodic synchrony.

2.4 General Discussion

Taken together, these experiments represent an important step towards establishing processing fluency as a potential factor in the perception of synchrony. Here we demonstrate that non-synchrony may reduce performance on a Flanker task, particularly when there are distracting stimuli that move faster than the target, but in the same direction. This is likely because fast motion may more easily capture our attention (Nothdurft, 2002), making those flanker stimuli more distracting. If all the stimuli are moving together at a faster rate, then the ability to perceive the target appears to be enhanced.

When the target stimulus moved in the opposite direction to the flankers, this appeared to decrease accuracy in some participants, as observed in both experiments. The fact that these outliers were all observed in the same condition is worthy of note, however this was a small number of participants and they were removed from analysis as statistical outliers. Some participants also self-reported that direction-incongruency felt particularly difficult, although this subjective feeling does not seem to have been reflected in their performance. Furthermore, one participant reported perceiving the stimuli as rotating, rather than moving, and further

investigation would be required to understand how the stimuli gave rise to this perceptual experience. Prior research has used scrolling vertical lines to induce a sense of rotation in participants (Zacharias & Young, 1981). However, the static Gabor patches were included in this study precisely to avoid such occurrences. Lange-Malecki and Treue (2012) did not report any individual differences within their sample, although they did find decreased performance on average for direction-incongruency using a similar animated Flanker design to that presented here.

The stimuli used here are very abstract, compared with most of the literature studying synchrony in social contexts (Mogan et al., 2017; Rennung & Göritz, 2016; Vicaria & Dickens, 2016). This task was chosen because it provides continuous movement which may be in synchrony without being rhythmic, as we were concerned that a strongly rhythmic stimulus may encourage participants to entrain their responses to the stimulus (Killingly et al., 2021), thus influencing reaction times and confounding the results. The animated Gabor patch provides a solution, as it has a continuous movement without providing a 'beat' that participants could tap in time with. Although this makes it difficult to apply the findings from the present study to real-world social contexts, it does establish a precedent for studying reaction times in synchrony perception studies.

Two dimensions of coordination were used in the present study. Direction-congruency could be thought of 'behaviour matching', while speed-congruency could be 'synchrony' using Bernieri and Rosenthal's (1991) terminology. Behaviour matching (imitation) occurs when actions are matched in kind, while synchrony involves actions to occur simultaneously in time (Bernieri & Rosenthal, 1991), although many studies involve behavioural synchrony in which actions are matched in both time and kind. However, this may be overly simplistic as, in naturalistic interpersonal coordination, movements often contain periodic reversals (rapid changes in acceleration) associated with the beat or tactus (Burger et al., 2014). Non-synchrony

in social contexts could contain differences in both speed and direction at varying times, although synchrony should involve speed and direction being matched at every timepoint.

The stimuli used in this study may be best understood as derivatives of periodic movement. They are continuous vectors, moving at a consistent velocity. As such, the conditions used in Study 1b may be plotted as linear functions (Figure 2.9).

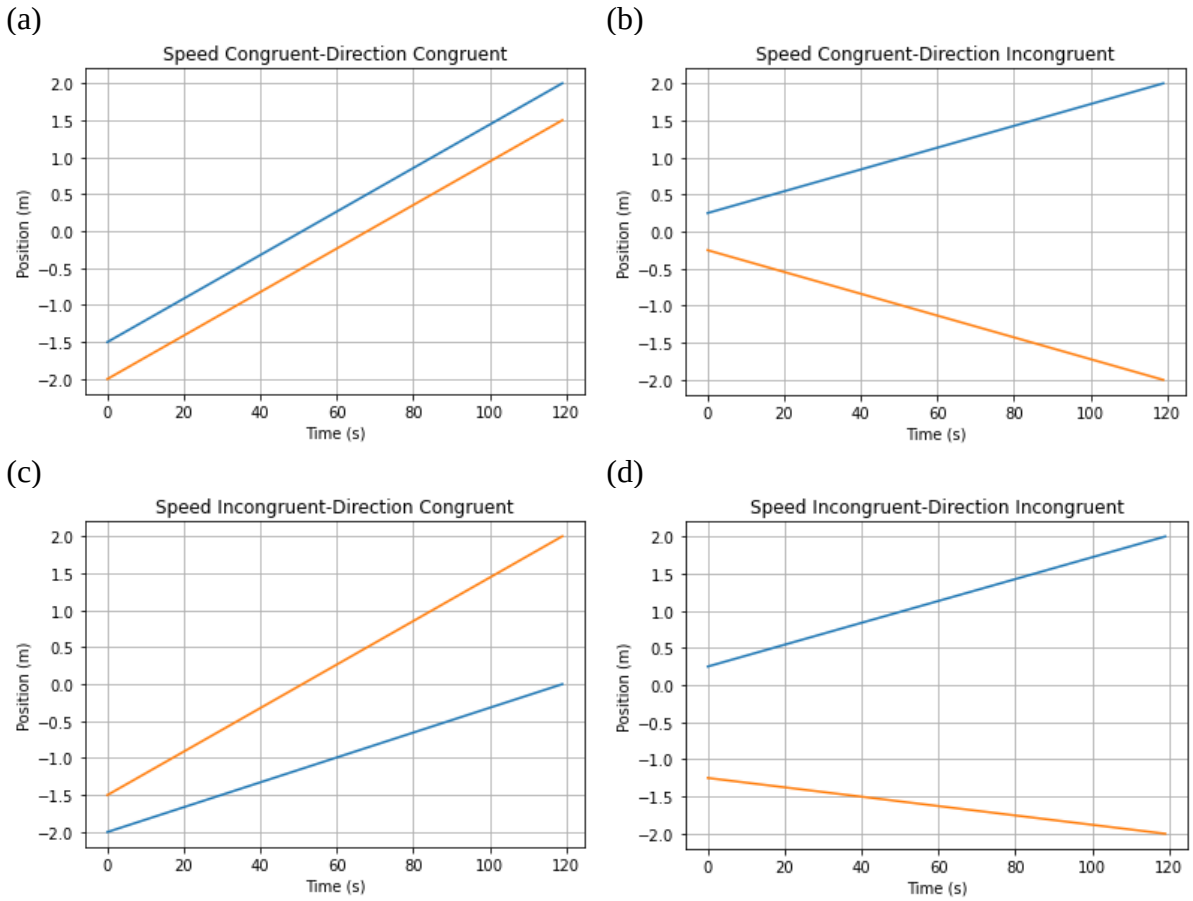
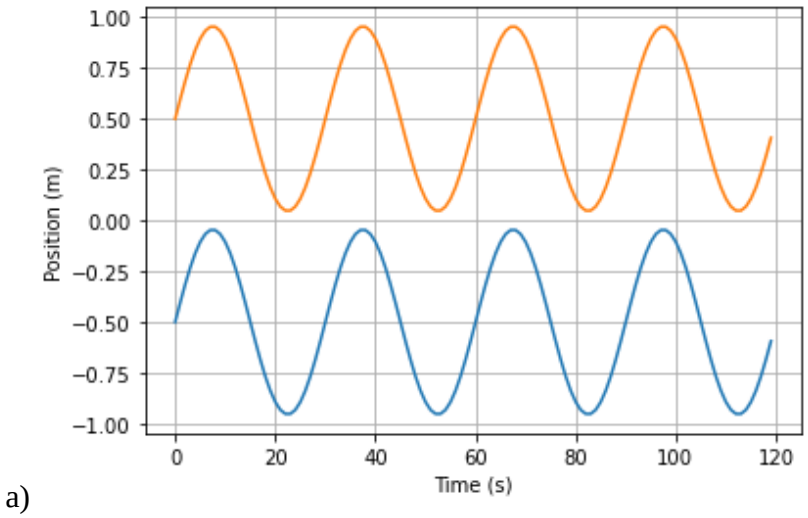
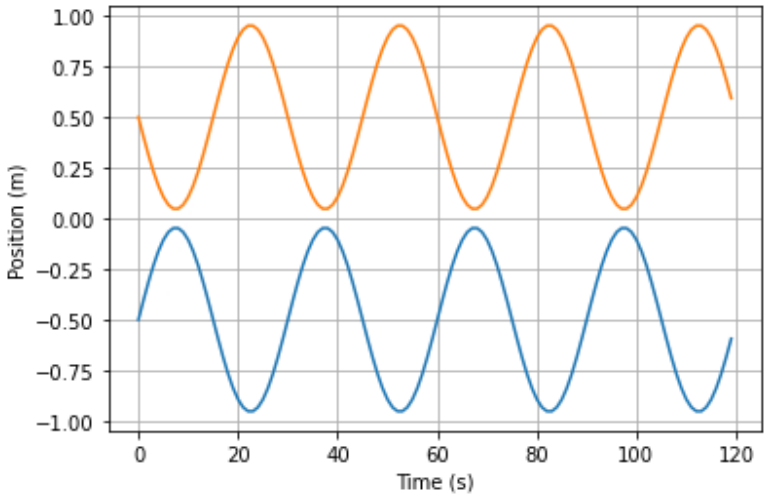


Figure 2.9 The conditions from Study 1b (blue representing the target and orange representing the flankers), plotted as linear functions showing change in position over time.

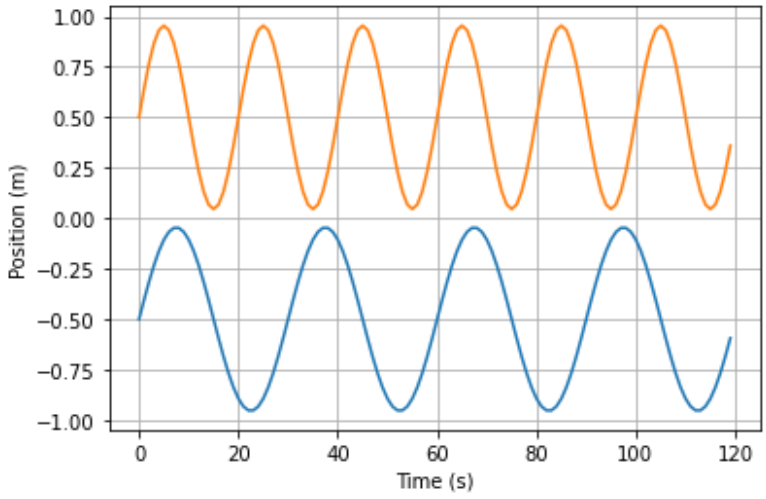
Synchronised, periodic movement emerges when two oscillators share the same frequency (Repp, 2005), such as in Figure 2.10a. The derivative of any timepoint in Figure 2.10a will always result in two parallel lines, as in Figure 2.9a. Therefore, the experience of perceiving the SC-DC condition resembles the experience of perceiving synchronised, periodic movement for a brief moment.



a)



b)



c)

Figure 2.10 Illustration of two oscillators (one in blue, the other in orange) moving over time in: (a) synchrony, (b) anti-phase synchrony, (c) non-synchrony.

However, two sine waves that are offset by half their period may still be considered synchronised, i.e., in an anti-phase relationship (Repp, 2005). A common example of this would be two hands clapping. Although the hands will always be moving in opposite directions, they would, nevertheless, be synchronised. Such a relationship may be seen in Figure 2.10b. The derivatives of these waves at any timepoint would always result in two vectors of equal but opposite slopes (except at the peaks and troughs), such as seen in Figure 2.9b, and corresponding to the SC-DI condition. Therefore, direction incongruency, without speed incongruency, is likely to emerge in the moment-to-moment perception of anti-phase synchrony.

Truly non-synchronised stimuli will rarely have speed congruency. The derivative of the waves in Figure 2.10c will, at any given timepoint, likely resemble the vectors in Figure 2.9c or Figure 2.9d – corresponding to the speed incongruent conditions. The only exception would be rare moments when the peaks and troughs may align. Therefore, the speed incongruent conditions may best capture the experience of a single moment of non-synchrony in periodic movement.

The stimuli we have used here may be best understood as derivatives of synchronised movement. The reduced performance observed in speed-incongruent (SI) trials in Study 1b indicates a higher processing load. If the SI condition is best understood as a snapshot of non-synchrony (Figure 2.10c), then this suggests higher processing load when observing non-synchrony. Conversely, SC conditions arise at every time-point during synchronised interaction (whether in-phase or anti-phase).

We suggest that speed congruency – a feature of periodic synchrony – is easier to process. Previous studies have found that Gabor patches moving at the same rate are more likely to be grouped together (Rideaux et al., 2016). The ease of processing observed here may

arise from this grouping effect. This finding provides an extension of our understanding of the perceptual experience of synchrony at single time-points, rather than just as a continuous experience.

Most studies into the synchrony-bonding effect have involved participatory synchrony. Our participants were not required to move or tap in time with the stimuli, only to respond by categorising stimuli. This makes it difficult to compare to previous work, like that of Hove and Risen (2009), but it did remove potential confounds that could arise from measuring reaction times while participants perform rhythmic actions (Killingly et al., 2021). Nevertheless, some previous studies have found that perception of synchrony can lead to a greater perception of beauty and more enjoyment, when observing others performing joint action (McEllin et al., 2020). This is consistent with the effects one would expect if viewers were experiencing processing fluency from a stimulus (Reber et al., 1998). The results here suggest that further research into processing fluency effects arising from the perception of synchrony would be warranted.

Future research is required to investigate the role of processing fluency in synchrony perception. Although our results suggest that non-synchrony may be more difficult to perceive than synchrony, future studies could target different stages in perceptual processing and executive function to examine how synchrony is processed. New tasks may need to be developed to target different stages of processing, or neuroimaging could be employed to determine which brain regions are most active during the perception of synchrony or non-synchrony. Furthermore, rhythmic stimuli in both the visual and auditory domains should also be studied, as most naturalistic synchronised movement involves rhythmic entrainment, as well as sensorimotor synchronisation across multiple sensory modalities (McNeill, 1997), rather than continuous visual motion as presented here. Finally, although this study suggests synchrony is easier to perceive than non-synchrony, it did not test whether synchrony is easier

to participate in, nor did it test for any affective or social consequences of the synchrony. To establish processing fluency as a possible factor in the synchrony-bonding effect, further investigation is required.

Most of the theories attempting to account for the social bonding effects of synchrony rely upon higher-order social cognition (Mogan et al., 2017). However, this study suggests a role of early-stage perceptual processing in the synchrony-bonding effect, at least for the visual domain. In particular, it suggests that synchronised movement may be easier to process than non-synchronised movement, and that this may occur in non-social stimuli. Most synchrony-bonding studies investigate interpersonal synchrony, and the low-level perceptual processes involved have often been ignored. This does not preclude other levels of explanation, and there may be interactions between different levels of cognition. Other contextual factors may influence the way synchrony is perceived or motivated (Miles, Griffiths, et al., 2010). This study only suggests that synchrony is easier to perceive, but how that ease of perception is interpreted may rely upon other factors.

The present study represents a first attempt to apply the processing fluency theory to the perception of synchronised action in a non-social setting and there is much scope to expand upon these findings. We tested the novel hypothesis that synchrony is easier to perceive than non-synchrony. We suggest that an experience of processing fluency may arise from the ease of organising (or grouping) synchronous stimuli. Based upon the observation that performance is most affected by distractors that move faster than the target, the disfluency may also occur from attention control. The stimuli presented here may be best understood as ‘snapshots’ of synchrony, as they are continuous vectors without the periodic movement usually associated with synchronisation between oscillators. As such, we demonstrate that when objects move at the same speed (but not necessarily in the same direction), they may be easier to process. This state of speed-congruency is a feature of both in-phase and anti-phase synchrony, but rarely

occurs when oscillators have different frequencies. While the present study does suggest a role of processing fluency in the perception of synchronised action, it does not demonstrate that the same would exist when actively participating in synchrony, or that the experience of processing fluency would necessarily lead to positive social evaluations. Further study would be required to gain a better understanding of the role that low-level perceptual processing may take in the synchrony-bonding effect.

2.5 Acknowledgements

This project received funding from the John Fell Fund. The authors wish to acknowledge Eric Clarke and Bahar Tunçgenç for their helpful comments on a draft of this manuscript.

3 Synchrony reduces cognitive load during a continuous tapping task

Joshua S. Bamford[1], Bronwyn Tarr[1,2], Lynden Miles[3], Emma Cohen[1]

[1]University of Oxford

[2]The London Interdisciplinary School

[3]University of Western Australia

3.1 Introduction

Processing fluency is a general principle that has been used to explain cognitive-affective preferences across a wide range of domains. The experience of processing fluency arises when a stimulus is relatively easy or quick to process – an experience which is 'hedonically marked' or rewarding (Winkielman et al., 2003). These rewards are then be associated with the stimulus, leading to increased perceptions of its beauty, truth, trustworthiness and liking (Reber et al., 2004; Topolinski, 2013). For instance, people will rate an image as more aesthetically pleasing if it is preceded by a matched prime, making the stimulus easier to perceive and identify (Reber et al., 1998). These effects also apply to social stimuli, as a people are more likely to give favourable ratings to others who are easier to process – e.g. difficulty in categorising others who do not fit within one's prototypes then leads to prejudice (Lick & Johnson, 2015). Although applied to many aesthetic and social experiences, processing fluency has never been explicitly tested in interpersonal synchrony.

Synchronisation involves the coordination of multiple agents in time, such that their actions are in temporal alignment, usually with a regular, periodic structure (Bernieri & Rosenthal, 1991). Often participation in synchronised action leads to subsequent feelings of social connection or bondedness, referred to as the synchrony-bonding effect (Hove & Risen, 2009; Mogan et al., 2017; Rennung & Göritz, 2016; Vicaria & Dickens, 2016). However, the precise cognitive mechanisms behind the synchrony-bonding effect remain unclear.

One suggestion is that interpersonal synchronisation may lead to a processing fluency effect, which in turn may lead to a social bonding effect (Wheatley & Sievers, 2016; see Chapter 1). The possibility that synchrony may be easier to process was briefly explored in a footnote by Hove and Risen (2009), but not investigated further. This may feasibly occur as synchrony is a cue for grouping visual (Rideaux et al., 2016) and auditory stimuli (Bregman, 1990; Hukin & Darwin, 1995). The reduction in quantity of information streams may then decrease cognitive load, as more cognitive resources are required to attend to more objects (Alnæs et al., 2014; Hughes et al., 2013; Zeamer & Fox Tree, 2013). Previous research has suggested that synchrony may be easier to perceive than non-synchrony, possibly because synchrony groups objects together (Lange-Malecki & Treue, 2012; see Chapter 2). These studies involved purely perceptual tasks, however, and did not require participants to engage in synchronised action.

It should be noted that maintaining interpersonal synchrony is a complex attentional skill. In music ensembles, for example, one must split attention between one's own actions and those of others in the ensemble (Keller, 2001, 2013). This is especially pronounced in DJs, for example, who must beat match between two different auditory streams, maintaining a mental representation of both (Witek, 2022). Thus, although maintaining synchrony requires dedicated attentional resources, synchronised action may still alleviate processing load when compared to non-synchronised action. However, this remains untested. The present study aims to investigate the difference in processing load during tasks with different degrees of synchrony. It uses a novel dual task paradigm in which participants complete a tapping task while also performing a secondary attention task. Performance on the secondary task is then used to measure the processing load of the tapping task.

3.1.1 Dual tasks and auditory distraction

Cognitive load research regularly employs dual-task methodology to assess the difficulty of a primary task by measuring performance on a secondary task (Abernethy, 1988). This is based on the idea of performance trade-offs: assuming there are limited cognitive resources available, if the primary task requires more attention, then there will be less available for the secondary task, thus sacrificing performance on the secondary task (Huang & Mercer, 2001). Although dual-task paradigms have never been implemented to study the processing load of synchrony, useful comparisons can be drawn from some of the general literature on auditory distraction.

Previous studies have looked at the performance impacts of auditory distraction. Distractions draw attention away from the primary task, disrupting performance on the task, but also potentially providing useful environmental information (Morgan et al., 2020); it is important to be able to shift attention towards new stimuli if they indicate a possible threat or opportunity. For example, so-called background music disrupts attention – particularly music with lyrics (Shih et al., 2012). Auditory stimuli that are more complex, with multiple melodic lines, are also more disruptive to language processing than simpler melodies (Fiveash et al., 2018)(Fiveash et al., 2018). In general, studies of distraction effects find that more stimuli require more processing to maintain attentional filters (Hughes et al., 2013; Zeamer & Fox Tree, 2013), and increasing the number of objects to attend to increases processing load (Alnæs et al., 2014).

Dual task paradigms are common for studying processing load. However, they have not yet been deployed to study the processing load associated with a synchronised tapping task. One study found that tapping performance could be affected by the difficulty of a concurrent memory task (Pecenka et al., 2013). In this case, the memory task was the primary task, and

tapping treated as the secondary task. The present study reverses these roles: placing the tapping task as the primary task.

3.1.2 Pupil measures of processing load

While many studies rely on behavioural measurements to study processing load, physiological cues may also be used. Although pupils may dilate for many reasons (Mathôt, 2018), Beatty (1982) first noticed that perception of a stimulus was often associated with a subsequent pupil dilation, and that the magnitude of this pupil dilation event was related to the processing load required by the stimulus. Pupil activity is tightly linked to activity in the locus coeruleus, responsible for regulating psychological arousal and autonomic activity (Samuels & Szabadi, 2008). Task-evoked pupil dilation is now well-recognised as a measure of processing load (Reilly et al., 2019). Although mostly studied with visual stimuli, pupils will also dilate to purely auditory stimuli (Bianco et al., 2019; Zekveld et al., 2018). Pupil activity also seems to indicate attentional load in non-human primates (Hampson et al., 2010). However, most pupillometry studies measure short windows of pupil activity immediately following stimulus presentation, which is not suitable for a continuous activity such as finger tapping.

In continuous tasks (i.e. without discrete events), there are other methods of assessing processing load. Mean pupil size may be a useful measure of load during longer, continuous tasks (Kiefer et al., 2016). Meghanathan and others (2015) found that both mean pupil size and mean fixation duration indicate load during a visual search task. As well as the average length of fixation, the number of fixations – referring to the number of times the eyes focus on a stimulus – may also indicate processing load (Jacob & Karn, 2003). More sophisticated analyses using wavelet transforms may also be used, although these are less well established in the literature (Duchowski et al., 2018). Rather than measuring pupil responses to specific stimuli, or taking an average pupil dilation, the wavelet transform extracts information about short term fluctuations in pupil size, which may indicate attentional processes (Duchowski et

al., 2018). It is thought that the ratio of high frequency to low frequency wavelets may be the most reliable measure of processing load, which may be measured by the Low/High Index of Pupillary Activity (LHIPA; Duchowski et al., 2020).

Pupillometry measures of processing load are starting to be deployed in social coordination research. A study by Bishop and others (2021) found that the technical difficulty of a piece of music increased average pupil size of the players in a string quartet. This may have been related to specific features of the music, which could include rhythmic complexity, however the degree of synchrony between players was not measured specifically. It remains to be tested whether synchrony may reduce pupil size compared with non-synchrony on an interpersonal coordination task.

3.1.3 Fluency through synchrony

The present study aimed to test whether synchrony is easier to process than non-synchrony during an interpersonal coordination task. To do so, we employed a dual task paradigm. The primary task in this study is based upon previous finger-tapping studies (Hove & Risen, 2009), although in this task participants were presented with a pre-recorded video of two people tapping on a table (Figure 3.1). Participants were instructed to tap in time with one of these people. Each of the video stimuli was created from two separate videos, stitched together to create the different conditions: one in synchrony, another in non-synchrony. In addition, a control condition was added, in which only the target person was tapping, and the other person (the distractor) in the video remained still. This was to establish whether the effects of synchrony could be attributed to perceptually grouping the stimuli together; the streaming account would suggest that there should be no difference in processing load between one stream (the control condition) and two streams (the synchrony condition).

A secondary task was employed as a measure of processing load. In this case, participants were asked to perform a visual search task, in which they had to correctly identify a target letter in amongst a range of distractor letters on screen (Treisman, 1977). The visual search task was chosen as each trial is expected to take longer than the period of the beat in the tapping task (0.5s at 120 beats-per-minute). This was to prevent participants from entraining their response to the visual search task to the beat of the tapping task, as previous research has found that participants will often synchronise responses between tasks when provided with a beat (Killingly et al., 2021). If the time taken to complete a trial of the secondary task were less than the period of the beat, then participants might delay their response in order to complete each trial on the beat, which would limit the usefulness of reaction time as a measure of task performance.

In a within-subjects design, each participant completed each condition (synchrony, non-synchrony, control) twice, for a total of six trials. In each trial, their mean reaction time (RT) and error rate (ER) on the secondary task were combined into a 'performance' score. Pupillometry was also employed as an additional measure of processing load. Since there is limited previous literature on pupillometry measures during synchronisation tasks, a range of different measures were used including the LHIPA, mean pupil size, and number of fixations.

Across the three conditions, we expected poorest performance on the visual search task during non-synchrony. If synchrony contributes to perceptual grouping, thus alleviating processing load by reducing the number of objects one must attend to, then we expect no difference between the synchronised and control conditions. We expected the same pattern of results with the pupillometry measures: relative to the other conditions, in the non-synchrony condition, mean pupil size should be largest, LHIPA should be lowest, and there should be the highest number of gaze fixations. Additionally, we collected tapping data to use as a control to

confirm that participants were not sacrificing primary task performance in favour of secondary task performance.

3.2 Method

This study was approved by CUREC at the University of Oxford (Ref: SAME_C1A_19_002) and the ethics board of the University of Western Australia (Ref: 00007908). The study design was preregistered.[10]

3.2.1.1 Participants

Participants were recruited from the student population at the University of Western Australia (See Appendix 6.2.1). There was a total of 104 participants (66 female, 35 male, 3 other), with a mean age of 23.6 (*SD* = 8.73, range = 18 to 61). Fourteen participants were left-handed. All participants reported having normal or corrected-to-normal vision, and that they normally use a computer without needing glasses. This was an important exclusion criterion, given that the eye-tracking headset cannot be worn with glasses. All participants gave informed consent (see Appendices 6.2.2 and 6.2.3) and were fully debriefed as to the aims of the study at its conclusion (see Appendix 6.2.5).

3.2.1.2 Materials

Video stimuli were all pre-recorded on a Sony RX-100 camera. Each video recording featured an adult male wearing a black t-shirt against a white background, tapping an isochronous beat on the table in front of him (see Figure 3.1). The man in the video was only visible from the neck down (i.e., torso and arms). Each video featured the same person to keep all aspects of the visual stimuli controlled, aside from the movement of the person's hand. As the participants were not asked any questions about the person in the video, it was not important

[10] https://aspredicted.org/7HQ_YLM

whether they believed him to be the same person each time. See Appendix 6.2.10 for example video stimuli.

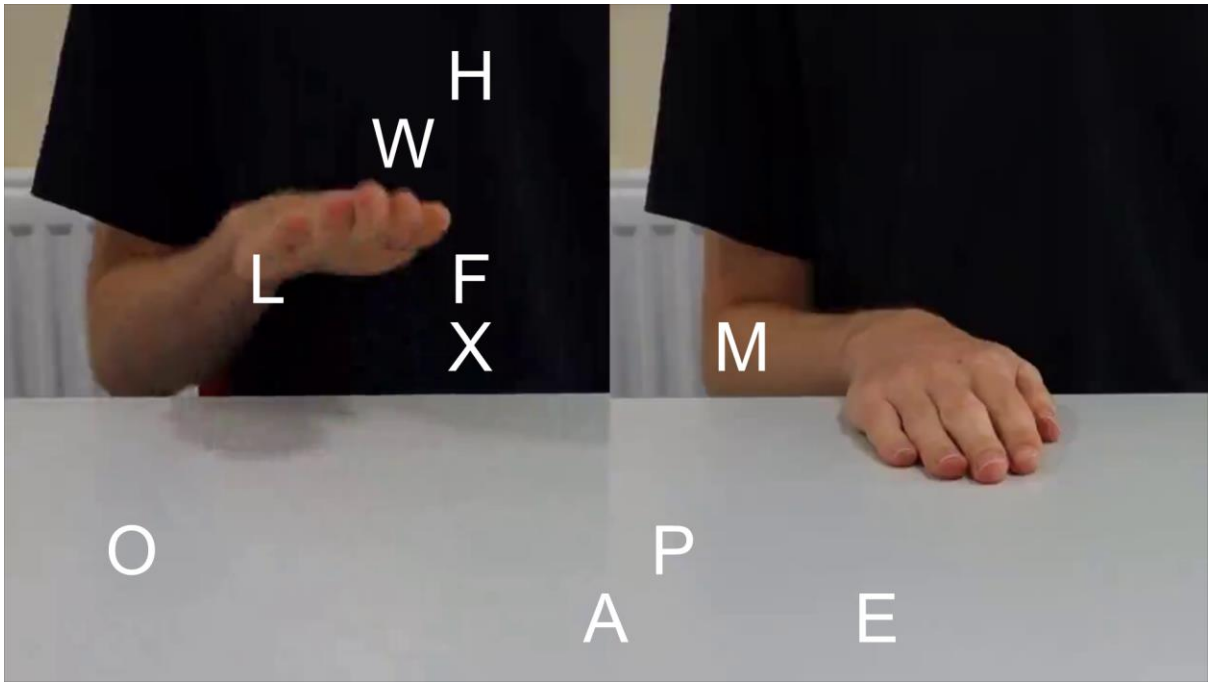


Figure 3.1 Screen capture during the experiment. The participant was asked to tap in time with the target on the left. In this instance, the distractor was in the still condition. Here a single trial of the letter search task appears superimposed on top of the video stimulus.

To create the videos, the man in the video was instructed to tap in time with a metronome. Three videos were recorded, each with a different metronome: 120 BPM, 110 BPM, and silent (no tapping). The 120 BPM metronome also introduced a shift in tempo of 1% approximately every 30 seconds. This alternated between speeding up, returning to baseline, then slowing down Because the 120 BPM video was to always be the target, the subtle variations in tempo would require the participant to maintain attention towards the target video, rather than relying upon an internal time-keeper.

The study stimuli were produced from these initial video recordings. Each stimulus comprised of two videos in a split-screen cut by the vertical axis. This was preceded by a four

second timer as a fixation point. Each stimulus had the 120 BPM video on the left, and either the 120 BPM, 110 BPM, or silent video on the right. This produced stimuli in three conditions: non-synchronised (Left:120 – Right:110), control (Left:120 – Right:Silent), and synchronised (Left:120 – Right:120).

PsychoPy v3.1.5 (Peirce et al., 2019) was used for stimulus presentation and for recording responses to the visual letter search task. This task was based upon those used by Treisman and Gelade (1980). In this task, an array of 9 letters appeared on screen: O, X, E, H, W, P, A, F, L. In addition to those letters was either an 'N' or an 'M', in equal proportions between trials. All letters were presented in a randomised location, and with a size defined as one tenth of the screen height. Participants were required to identify either the 'N' or 'M', without a time limit. There would then be a short pause of 50ms before the next presentation. These trials were presented in blocks of 50. The visual letter search trials were presented superimposed in a layer in front of a video stimulus (see Figure 3.1) in one of the tapping conditions (synchronised, non-synchronised, control), with one condition per block. Condition order was randomised within the script and order was included as a covariate in analysis. The code was initially generated using the PsychoPy builder and then manually edited, as the builder does not natively support multiple, concurrent tasks. Video stimuli were presented on a 27" monitor with 1920x1080 resolution and a refresh rate of 60Hz. The code for this experiment and video stimuli are available in the supplementary online materials.[11]

Participants heard the stimuli through standard headphones. The participants' tapping during the experiment was recorded using an Adeline AD-35 contact microphone attached to the desk. Both the audio output from the computer playing the stimuli, and from the contact microphone, were fed to a Zoom H4n audio recorder.

[11] https://github.com/Social-Body-Lab/fluency-through-synchrony/tree/main/Study2/Study2_Experiment

Eye movements were tracked using a Pupil Core eye-tracker from Pupil Labs (Kassner et al., 2014),[12] equipped with a single camera under the right eye of the participant. This was calibrated using the default calibration procedure within the Pupil Capture software, with QR markers to localise the monitor. Control of the software was done through the Pupil API within the PsychoPy script.

Individual differences in personality traits were captured using a standard 10-item measure of the big five (BFI-10; Rammstedt & John, 2007).[13] Musicality was measured using the Perceptual Ability and Musical Training subscales of the Goldsmiths Musical Sophistication Index (Gold-MSI; Mü Llensiefen et al., 2014).[14] These scales were only used for descriptive purposes in this study and were not analysed further. A demographic questionnaire was also included (see Appendix 6.2.8).

3.2.1.3 Procedure

Each participant was welcomed into the lab and briefed on the study. They were also asked to confirm that they were aware of the exclusion criteria (i.e. needing glasses to use a computer) and instructed to remove any watches, rings or bracelets on their left hand that may interfere with the audio recording. The Pupil Core headset was then fitted and calibrated within the Pupil Capture software. PsychoPy was then initiated, and the participant was left to guide themselves through the protocol. Throughout the session, the experimenter was sitting out-of-view on the opposite side of the room, and participants were made aware that they could ask any questions during the practice trials of the experiment.

The participant was presented with on-screen written instructions for the task (see Appendix 6.2.4). They were instructed to tap on the desk in front of them with their left hand,

[12] https://pupil-labs.com/products/core/
[13] See Appendix 6.2.6
[14] See Appendix 6.2.7

and to do this in time with the person on the left side of the video. While tapping with their left hand, they would be performing a letter search task with their right hand. Letters would appear on screen, and amongst them would be either an 'N' or an 'M'. The participant was asked to press the corresponding 'N' or 'M' key when the target letter was found. There was no time limit to find the target letter. They were then given three short blocks of practice trials in each of the three conditions, in randomised order (10 letter search trials in each block) and the video ended once the letter search trials in that block were complete. After the practice trials, participants were presented with the main task in six blocks. Each block was randomly assigned to one of the three conditions (synchrony, non-synchrony, control). The randomisation was coded to ensure each participant experienced each condition once in the first three blocks. See Figure *3.2* for an overview of the design. For every trial, the reaction time (RT) to find the target letter was recorded, as was whether the participant's response was correct.

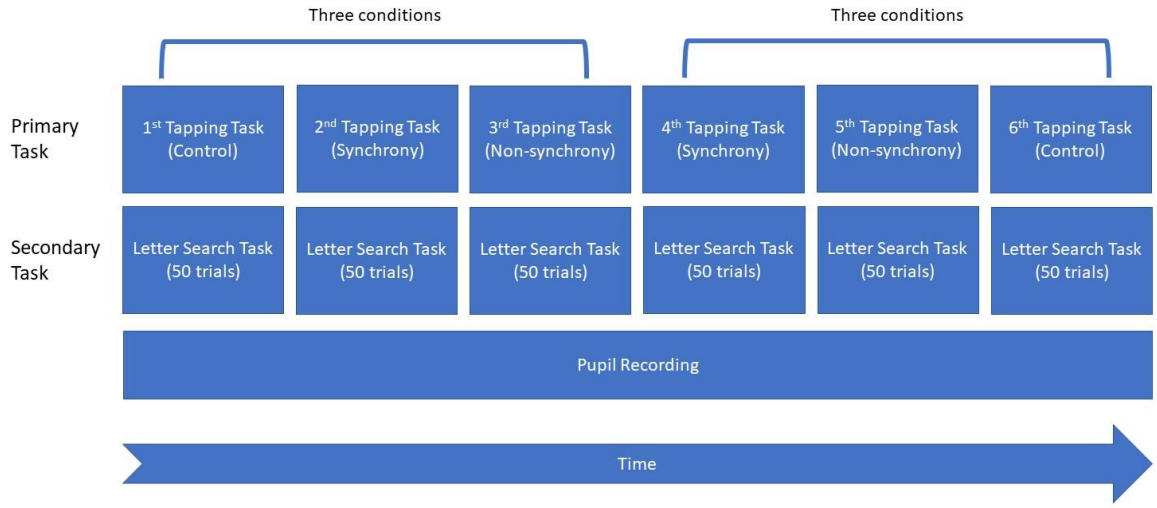


Figure 3.2 Experiment design showing the six blocks, each of 50 trials in the Letter Search Task, and with an example of the three conditions for the tapping task in randomised order, presented twice.

After completing all six blocks, the participant was fully debriefed as to the aims of the study (see Appendix 6.2.5). During the debrief the participant was asked what they thought the hypothesis was, and whether they found any of the conditions harder than any others, as a hypothesis probe.

3.2.1.4 Data Preparation and Analysis

Data were processed in Spyder using the Pandas library for Python. Data from the letter search task were imported from PsychoPy and combined with Pupil and tapping data. The mean reaction time (RT; seconds) and error rate (ER; 0-1, where 1 corresponds to no errors) for the letter search task were calculated per block of trials. These were used to determine a combined performance score (ER / RT) for each block of trials.

Pupil recordings for each participant were imported and segmented by block. Eye-blinks were detected by the Pupil Core Blink Detector using default settings and were subsequently removed from each segment. Data were additionally checked to ensure all frames without a measured pupil size were removed. Pupil size was determined by the 3D model produced by the Pupil Core (Świrski & Dodgson, 2013). This 3D model is more robust to changes in gaze direction, compared with regular 2D pupil size estimates, so there was no need to correct for the foreshortening effect (Petersch & Dierkes, 2021). The mean pupil size was then calculated across all frames per segment. The High/Low Index of Pupillary Activity was also determined for each segment. This takes the ratio of high to low frequency wavelets; the procedure and code for this measure is described by Duchowski and others (2020). Fixations were identified using the offline fixation detection plugin for Pupil Core, and these were segmented by block accordingly.

Audio recordings were imported into Spyder using the Librosa library (McFee et al., 2022). Audio was recorded in stereo, with the left channel recording the audio output from the

video stimuli, and the right channel recording the participant's tapping from a contact microphone. The left channel was used to segment the audio, marking the start and end times of each video stimulus. These segments were then applied to the right channel. The Librosa beat detection method extracted beats from the participant's tapping. Each absolute beat time was converted to an Inter-Onset Interval (IOIs) by subtracting the time of the previous beat. The mean and variance in IOI were calculated for each segment, as measures of tapping performance.

Once imported and processed, data were analysed using the Pingouin library (Vallat, 2018), and figures were produced in Seaborn (Waskom, 2021). Due to technical errors in both the pupillometry and audio recordings, there were some missing data. These datapoints were excluded from their relevant measures during analysis (details below), but there were otherwise no other exclusions. Data were analysed with 1-way repeated measures ANOVAs with tapping condition as a within-subjects factor with three levels (Synchrony, Non-synchrony, Control). A separate RM-ANOVA was conducted for each dependent variable, after checking the assumptions of a RM-ANOVA for that variable. IOI Variance and IOI Mean were the first to be tested, as these were used confirm that participants followed the instructions to prioritise accuracy of tapping. To test the primary hypothesis that non-synchrony would reduce performance on the visual letter search task, a RM-ANOVA was then conducted on the performance (ER/RT) scores. This was followed by RM-ANOVAs for the pupillometry measures: pupil size, number of fixations, and LHIPA. Post-hoc, Bonferroni corrected t-tests were then conducted between the three conditions for all dependent variables. T-tests were single-tailed when testing the directional hypothesis that synchrony and the control condition would have higher performance, lower pupil size, fewer fixations and a higher LHIPA, than non-synchrony. Two-tailed t-tests were used between synchrony and control, as no difference was expected. Because a small effect of tapping condition on IOI variance was found, an

additional Linear Mixed-Effects Model (LMEM) was constructed to investigate any possible interaction effects between IOI Variance and Condition on performance, while including participant ID as a random effect.

Data and code for analysis is available in the supplementary online materials.[15]

3.3 Results

Descriptive statistics for all three conditions (control, non-synchrony, synchrony) are given below for each of the dependent measures (see Table 3.1). Inter-Onset Interval (IOI) variance and mean IOI were measured to investigate any possible confounds from participants sacrificing tapping accuracy. Mean IOI should be approximately 0.5s, as this is the expected IOI for a tempo of 120 BPM (two beats per second). The performance measure was calculated as error rate over reaction time as a combined measure of task performance in the visual letter search task. As the use of pupillometry was exploratory in this study, three different pupillometry measures were recorded. This included pupil size, the Low/High Index of Pupil Activity (LHIPA), and number of fixations. In this instance, *n* refers to the total number of blocks (two blocks per condition, per participant). Performance had a complete dataset (208). However, the tapping measures were affected by some blocks in which the recording quality was too poor to reliably extract beats due to a faulty cable. Some data was also missing from the pupillometry measures due to a software error that prevented the eye-tracker from booting correctly.

Two 1-way repeated measures ANOVAs were used to investigate the effects of tapping condition on Mean IOI and IOI Variance, as any such effects could provide potential confounds in later analyses. Significant effects were observed for both Mean IOI, $F(2, 134) = 9.87$, $p < .001$, and IOI Variance, $F(2, 122) = 16.21$, $p < .001$, in Repeated Measures ANOVAs. For the

[15] https://github.com/Social-Body-Lab/fluency-through-synchrony/tree/main/Study2/Study2_Analysis

Mean IOI measure, post-hoc Bonferroni adjusted t-tests revealed a significant difference between Control and Non-synchrony, $t(67) = -4.20$, $p < .001$, $d = -0.46$, and between Synchrony and Non-synchrony, $t(67) = -3.39$, $p < .01$, $d = -0.39$, but not between Synchrony and Control (see Figure *3.3*). The observed Mean IOI was slightly below the expected 0.5s in all conditions. However, the results indicate that participants were anticipating the beat more in the Control and Synchrony conditions compared with the Non-synchrony condition. Within the IOI Variance measure, post-hoc Bonferroni adjusted t-tests found a significant difference between Control and Non-synchrony, $t(61) = -4.80$, $p < .001$, $d = -0.45$, and between Synchrony and Non-synchrony, $t(61) = -4.10$, $p < .001$, $d = -0.49$, but no difference between Synchrony and Control (see Figure *3.3*). This indicates that IOI Variance was significantly higher in the Non-synchrony condition compared with the other two conditions. Because of this difference in IOI Variance, an LMEM was used to investigate IOI Variance as a potential covariate in later analysis.

Table 3.1 *Descriptive statistics.*

	Condition	*n*	*M*	*SD*
Performance	Control	208	0.7512	0.1455
	Non-synchrony	208	0.6862	0.1459
	Synchrony	208	0.7582	0.1403
Pupil Size	Control	140	4.943	0.8172
	Non-synchrony	140	5.072	0.8276
	Synchrony	140	5.014	0.8340
LHIPA	Control	140	2.445	0.7000
	Non-synchrony	140	2.401	0.6645
	Synchrony	140	2.364	0.2550
Fixations	Control	140	475.4	81.50
	Non-synchrony	140	505.9	107.9
	Synchrony	140	470.3	79.58
IOI Variance	Control	121	0.007895	0.007744
	Non-synchrony	123	0.01096	0.009293
	Synchrony	124	0.007069	0.005203
Mean IOI	Control	128	0.4838	0.02974
	Non-synchrony	132	0.4985	0.03614
	Synchrony	133	0.4868	0.02786

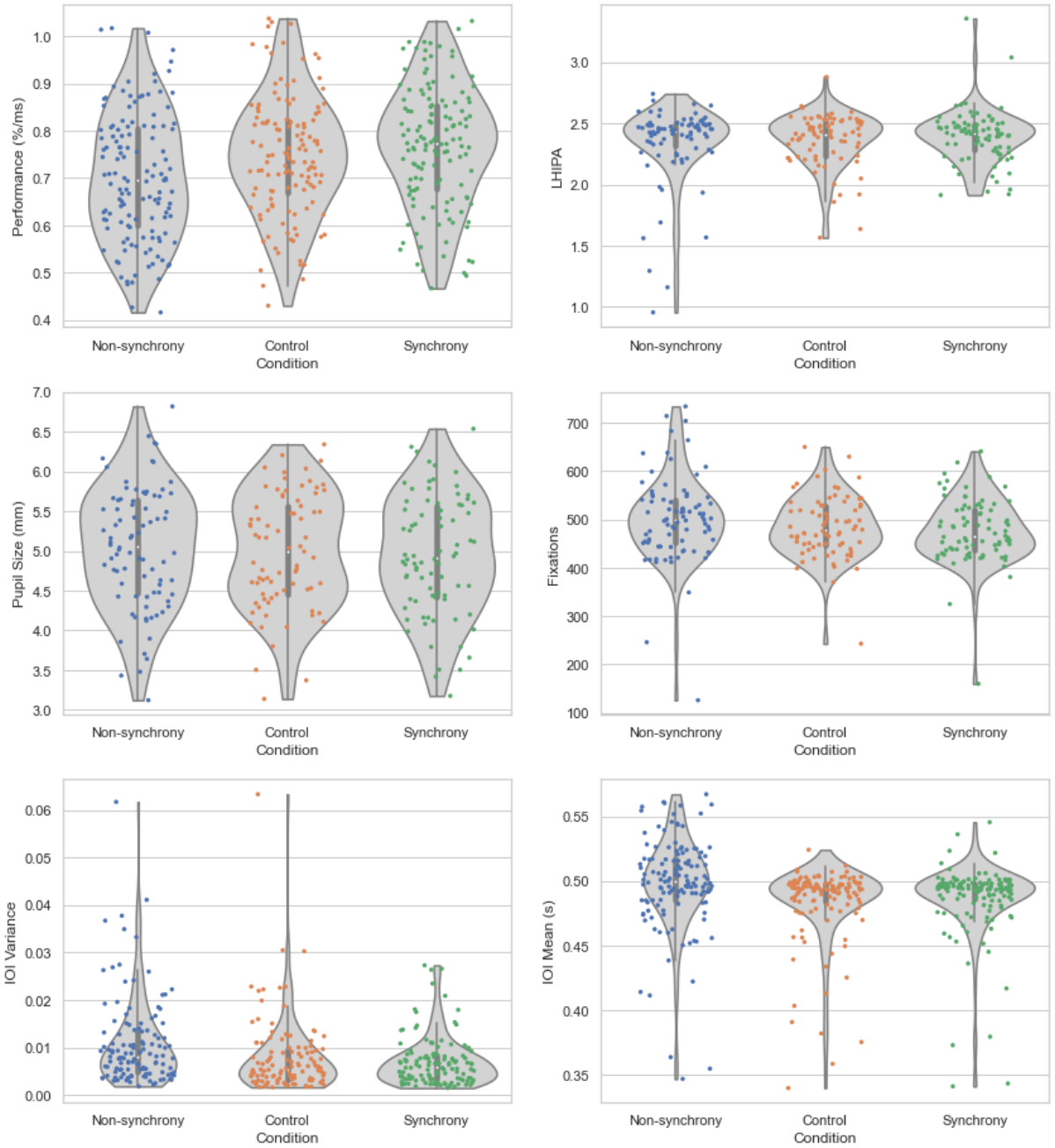


Figure 3.3 Violin plots for performance (ER/RT), LHIPA, averaged pupil size, number of fixations, IOI variance, and mean IOI by condition. Internal boxplot shows median, first quartile, and 1.5 quartile.

Having established the possible effects of condition on tapping accuracy, the effects of condition on visual search task performance were then analysed. The RM-ANOVA, testing the effects of the three levels of tapping condition on performance, violated the assumption of sphericity, $W = 0.877$, $p < .01$, so a Greenhouse-Geisser correction was applied. The corrected

RM-ANOVA found a large, significant effect of condition on performance, $F(2,206) = 68.63$, $p < .001$, $\eta_p^2 = 0.400$ (see Figure 3.3). This effect was maintained when controlling for block order, using a LMEM for the effect of condition on performance, with participant and block number included as random effects (see Appendix 6.2.9 for details). Post-hoc analysis revealed significant differences between control and non-synchrony, $t(103) = 8.68$, $p < .001$, $d = 0.467$, and between synchrony and non-synchrony, $t(103) = 10.00$, $p < .001$, $d = 0.528$, but no difference was observed between synchrony and control.

To further explore the effect of condition on tapping accuracy, an LMEM was constructed to predict performance from the interaction between IOI variance and condition, while controlling for participant as a random effect. The model found an effect of non-synchrony, and no effect of synchrony, relative to the control condition. An independent effect of IOI variance was found, but no interaction effects between beat variance and the tapping conditions on performance (see Table *3.2*). As seen in Figure 3.4, increase in IOI variance correlated with a decrease in performance across all conditions.

Table 3.2 *Linear mixed-effects model for performance by condition (compared to control) and IOI variance, controlling for participant.*

	Effect	Estimate	*SE*	*df*	*t*	*p*
	(Intercept)	0.767	0.017	105.4	44.00	<.001
	Non-synchrony	-0.051	0.011	298.0	-4.584	<.001
	Synchrony	0.020	0.012	299.1	1.730	.085
	IOI Variance	-2.408	0.906	355.5	-2.657	.008
Non-synchrony	IOI Variance	0.447	0.898	298.0	0.498	.619
Synchrony	IOI Variance	-0.825	1.275	301.0	-0.647	.518

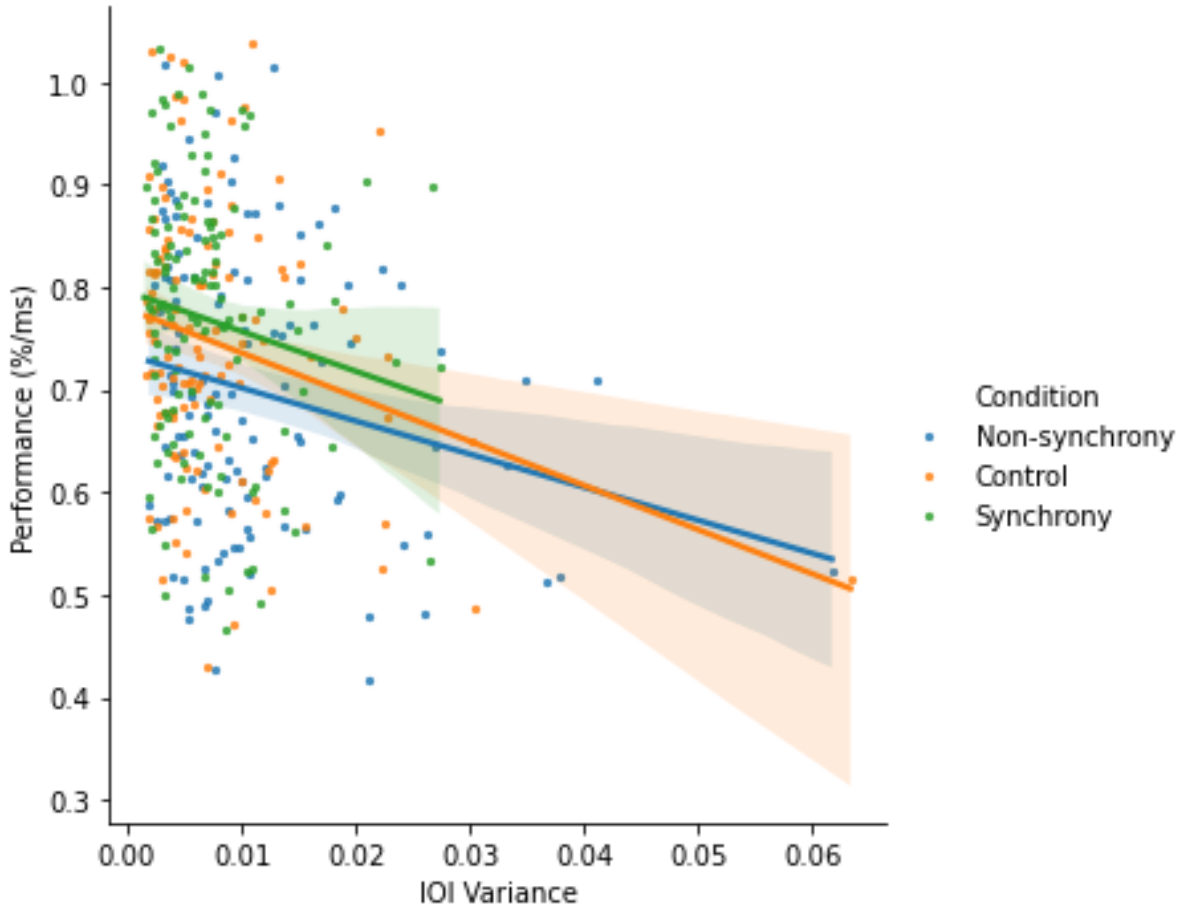


Figure 3.4 Linear mixed-effects model for performance by condition and IOI variance. The shaded area represents 95% confidence intervals around the regression line.

Pupil Size data met the assumptions for a 1-way Repeated Measures ANOVA, and a significant effect of condition was found, $F(2,138) = 11.85$, $p < .001$, $\eta_p^2 = 0.146$ (see Figure 3.3). Post-hoc, Bonferroni corrected, single-tailed t-tests found a significant difference between control and non-synchrony, $t(69) = -4.84$, $p < .001$, $d = -0.16$, and between synchrony and non-synchrony, $t(69) = -2.25$, $p < .05$, $d = -0.07$, with non-synchrony having larger mean pupil size in both instances. A two-tailed t-test found a significant difference between control and synchrony, $t(69) = -2.61$, $p < .05$, $d = -0.09$, with Synchrony having a slightly larger pupil size on average (see Table 3.1).

Number of Fixations and the Low/High Index of Pupillary Activity (LHIPA) were used as secondary eye-tracking measures of cognitive load. The data for fixations and LHIPA

violated the assumption of sphericity for a Repeated Measures ANOVA, and thus Greenhouse-Geisser corrections were applied to the p-values. A significant effect was found of condition on number of fixations, $F(2,138) = 26.03$, $p < .001$, $\eta_p^2 = 0.274$ (see Figure 3.3), however there was no significant effect on LHIPA, $F(2,138) = 1.23$, $p = .29$ (see Figure 3.3). Post-hoc, Bonferroni corrected t-tests found significant differences in fixations between control and non-synchrony, $t(69) = -4.971$, $p < .001$, $d = -0.34$, and between synchrony and non-synchrony $t(69) = -5.668$, $p < .001$, $d = -0.40$, with non-synchrony having more Fixations in both cases (see Table 3.1). There was no significant difference between control and synchrony on fixations, $t(69) = 1.792$, $p = .232$, $d = 0.06$.

Finally, during debriefing participants were also asked to verbally report which of the conditions they found the hardest. 90 out of 105 identified the non-synchrony condition as being the most difficult. The remaining 15 mostly reported no subjective difference in difficulty between conditions.

3.4 Discussion

This study aimed to test whether non-synchrony is more difficult to process than synchrony, during a coordinated tapping task. A control condition was also included, with just a single hand tapping, to test the hypothesis that performance enhancement during synchrony is due to stimulus streaming. Processing load was measured using performance on a secondary task and through multiple eye-tracking measures.

As predicted, performance on the secondary task was poorest in the non-synchrony condition, compared with both the synchrony and control conditions. This suggests that the non-synchrony condition made tapping in time with the target stimulus more difficult, as the non-target stimulus was tapping out of time, providing a distraction. Assuming that there are limited cognitive resources available (Huang & Mercer, 2001), the additional attention required

for the tapping task during Non-synchrony detracted from performance on the secondary task. This is consistent with the suggestion that Synchrony may be easier to process than Non-synchrony (Wheatley & Sievers, 2016). It appears, however, that there was no difference between Synchrony and Control conditions, consistent with the hypothesis that synchronised stimuli may be streamed together 'as one'.

Synchrony is a strong cue to both visual grouping (Rideaux et al., 2016) and auditory streaming (Bregman, 1990; Hukin & Darwin, 1995) resulting in multiple stimuli being perceived as one object. Perceiving multiple objects, or having to attend to multiple auditory streams, may be more difficult than a single auditory stream (Alnæs et al., 2014; Hughes et al., 2013; Zeamer & Fox Tree, 2013). If participants effectively perceived the Synchrony condition to have only one moving object due to streaming/grouping, then we would expect the processing load to be no different to the Control condition in which there was only one moving object. Indeed, there was no difference between Synchrony and Control on secondary task performance, suggesting that this perceptual grouping may be occurring. An alternative explanation is that the distractor hand was filtered out of perception as it provided redundant information when moving in synchrony with the target hand. This is plausible, although the present study is unable to distinguish between the grouping and redundancy explanations.

The effect of condition on the participants' tapping was also measured to account for possible confounds. As a dual-task design, this study assumes that there is a performance trade-off between the primary (tapping) and secondary (visual search) tasks (Abernethy, 1988). Increasing the difficulty of the primary task should reduce performance on the secondary task, as we observed. However, it is possible that participants may have also sacrificed performance on the primary task, in order to retain a higher standard of performance on the secondary task (Huang & Mercer, 2001). Additionally, under high load, participants may become more effective at filtering out task-irrelevant stimuli (Theeuwes et al., 2004). This was mitigated by

instructing participants to prioritise keeping in time with their target beat, and through slightly shifting the tempo of the target beat by 1% at irregular intervals. This tempo shift was intended to keep attention directed towards the synchrony task, while not being so great a shift to be disruptive. Some participants did report noticing the tempo shift when it occurred, and that it did momentarily increase the difficulty of the task. However, the presence of the tempo shift was held constant across conditions so we do not expect this to have introduced any confounds. Additionally, we also measured Mean IOI (Inter-Onset Interval) and IOI Variance, to investigate primary task performance. Ideally, there should be no difference of Mean IOI or IOI Variance between conditions, as this would indicate that participants were following instructions and prioritising the tapping task. However, we observed significantly more IOI variance in the Non-synchrony condition, suggesting that participants were maintaining a less stable beat while tapping in this condition compared with Synchrony or Control. In addition, the Mean IOI in Non-synchrony was different to the other conditions. Participants in the Non-synchrony condition tapped, on average, every 0.498 seconds. Given the beat of the target stimulus was at 120 BPM, this is very close to the expected period of 0.5 seconds between beats. However, the mean IOI for both the synchrony (M = 0.487s, SD = 0.0278) and control (M = 0.484s, SD = 0.0297) conditions was slightly below this expected period, indicating anticipation of the beat. People often anticipate an isochronous beat, leading to a 'negative mean asynchrony' or a slightly shorter IOI than expected (Repp, 2005). However, if one maintains a shorter IOI than the target beat then one would progressively get out of synchrony, so some corrective actions may be required. It is also possible that participants in this study may have either added extra beats or missed beats, which would have shifted their mean IOI away from the expected 0.5 seconds. Nevertheless, the anticipatory mechanisms required to maintain a negative mean asynchrony may have been disrupted in the non-synchrony condition putting the participant into a relatively more reactive rather than anticipatory mode of tapping.

Taking the mean IOI and IOI variance measures together, it seems performance on the tapping task was negatively impacted by the non-synchrony condition. Participants may have found it harder to maintain a consistent beat, or to anticipate the beat. This could be because the non-target stimulus was more distracting when tapping out of time, obscuring the target beat from attention. If this increase in IOI variance was also associated with an increase in performance on the secondary task, this would indicate a possible confound in the results, however this was not the case. However, significant effects were still found on secondary task performance. When statistically accounting for IOI variance within a linear mixed-effects model, the effect of condition on performance persists, with performance being lowest in the non-synchrony condition. Across all conditions, higher IOI variance was associated with lower performance on the letter search task; this suggests an additive effect whereby participants who had more difficulty staying in time with the tapping task also performed worse on the letter search task. These measures of tapping performance would seem to reinforce the finding that non-synchrony was more difficult to process than synchrony and indicate that secondary task performance was sufficient to measure this effect.

Additional measures of processing load were recorded using eye-tracking and pupillometry, as physiological measures of processing load. Pupil size was larger in the non-synchrony condition than synchrony, which was larger than the control. The task-evoked pupillary response is usually larger for a stimulus that is more difficult to perceive (Beatty, 1982). In long trials, which may contain many task-evoked pupillary responses, previous research has used a mean pupil diameter to assess processing load (Bishop et al., 2021; Kiefer et al., 2016), as we have done here. The larger mean pupil size during non-synchrony trials relative to the synchrony and control trials indicates greater processing load. However, whereas there was no difference observed between synchrony and control in task performance, mean pupil size was greater during synchrony than in control.

Even if the stimuli are being perceptually grouped together – reducing processing load – there may have been other factors influencing pupil size, unrelated to attention control and processing load (Mathôt, 2018). For instance, there was a greater quantity of movement on the screen during synchrony compared with control. Previous research has found that stimulus motion can induce greater pupil dilation (Wang et al., 2021), which could have been a possible confound for the control condition in this study. This highlights the difficulties with using pupillometry measures with naturalistic stimuli, and the importance of using multiple measures.

An increased number of fixations during a task is commonly associated with higher processing load. This is particularly relevant during a visual search task, as more fixations indicate a less efficient search experience (Jacob & Karn, 2003). In our study we indeed found a higher number of fixations in the non-synchrony condition compared with synchrony and control. Just as with the performance data, there was no difference between synchrony and control, reinforcing the idea that two hands tapping may be perceptually grouped together, and thus provide little additional perceptual load over one hand tapping. This is not consistent with the pupil size results, although pupil size may have been more susceptible to differences in the quantity of movement, as described above, while fixations may have been more associated with the difficulty in controlling attention when performing the visual search task.

No significant difference in the Low/High Index of Pupillary Activity (LHIPA) was detected between conditions. The LHIPA measure is a more exploratory measure of processing load (Duchowski et al., 2020). It analyses the ratio of low to high frequency wavelets over a course of a pupillometry recording. A lower LHIPA score indicates greater processing load. As a novel measure of pupil activity, it has only been tested on a limited range of tasks and conditions. It was included here as using a wavelet transform to analyse pupil activity should be more resilient to the effects of light levels and the foreshortening effect than absolute

measures of pupil size (Duchowski et al., 2018). The original Index of Pupillary Activity (IPA) was found to be susceptible to distortion related to gaze direction (the foreshortening effect), which would be problematic in a visual search task, so the LHIPA (Duchowski et al., 2020) was implemented in the present study. Initial validation of the LHIPA found that it may lack the sensitivity to detect small differences in cognitive load, as in one of the experiments it could only distinguish between the task and control, but not between different levels of difficulty within the task (Duchowski et al., 2020). Although we did not collect subjective self-report data on perceived difficulty, many participants anecdotally reported that they found the dual-task nature of this experiment to be very challenging. Thus, the LHIPA may have been affected by a ceiling effect, preventing it from distinguishing between the different levels of load between conditions. Although the LHIPA has been tested with an eye-typing task – very similar to the letter search task – it is yet to be validated with tapping tasks (Duchowski et al., 2020). The rhythmic nature of the tapping task may have also influenced pupil activity, as each beat would be a stimulus that could cause a pupil dilation event (Zekveld et al., 2018). Mean pupil size may have been robust to this effect, while a time series analysis such as the LHIPA may have been more sensitive to the rhythm inherent in the stimuli. More research may be required to determine whether the LHIPA can accurately assess cognitive load in a finger tapping task.

The present study may have been limited in its ecological validity. There are many natural music-making situations which may involve split attention, in which one must coordinate between multiple people playing separate parts in a music ensemble or when a DJ has to beat-match two different audio tracks (Keller, 2001; Witek, 2022). However, it is unusual to perform a visual search task during synchronised activity. During the debriefing process, one participant in the present study with significant percussion training stated that this task may be useful for percussion students to practise focusing their attention. Aside from this

possible pedagogical application, the task remains limited in terms of its generalisability to ordinary musical practice.

3.5 Conclusion

Taken together, the measures of performance, pupil size and fixations all indicate a higher cognitive load of the non-synchronised video stimuli presented in this study, when compared with synchronised and control stimuli. Furthermore, there appeared to be no difference between synchronised and control conditions. This is consistent with the theory that synchronised stimuli may be streamed together, and thus are as fluent to process as a comparable single stimulus. The correspondence between these measures mean that eye-tracking measures may be sufficient to infer cognitive load in the absence of any behavioural measurement based on task performance, offering a possible way forward for testing the fluency account in tapping tasks that involve social appraisal (rather than performance) measures.

Future studies should aim to replicate these results, but without relying upon a dual-task paradigm to assess cognitive load. The secondary task necessarily distracts from the experience of synchrony, meaning that we could not reliably ask participants about any social judgements or outcomes in this experiment. The paradigm presented provides initial evidence for increased processing load during non-synchrony, and may help to inform future research to investigate attention and processing load during synchronised action.

3.6 Acknowledgements

Thank you to the University of Western Australia for the use of their facilities, to members of the St Peter's College Choir for pilot testing, and to Laura Bishop for her helpful comments regarding pupillometry. The authors would like to acknowledge that the data for this

study was collected on Whadjuk Ngoonar land. This project received funding from the John Fell Fund.

4 Synchrony-bonding effect mediated by perceived difficulty

Joshua S. Bamford[1], Bronwyn Tarr[1,2], Lynden Miles[3], Emma Cohen[1]

[1]University of Oxford

[2]The London Interdisciplinary School

[3]University of Western Australia

4.1 Introduction

The synchrony-bonding effect has often been observed but the mechanisms behind it remain poorly understood. An association between synchronised movement and social bonding has been found in a wide variety of activities, such as singing, dancing, marching, and many sports (McNeill, 1997). Results from controlled experiments have suggested that even simple synchronised finger tapping can lead to social bonding effects within a dyad (Hove & Risen, 2009), and that this effect can scale for larger groups engaging in synchronised behaviour (Launay et al., 2016; Weinstein et al., 2016). These studies conclude that motor-based synchrony may be the active ingredient in the social bonding effects of coordinated group action. Reviews of the synchrony-bonding literature find a small but significant effect across these experiments (Mogan et al., 2017; Rennung & Göritz, 2016; Vicaria & Dickens, 2016). Theorists have suggested a wide range of mechanisms that may contribute to this social bonding effect (Kirschner & Tomasello, 2009; Valdesolo & DeSteno, 2011; Wolf et al., 2016). Most of these proposed mechanisms rely on higher-level cognitive appraisal or Theory of Mind. Here we propose a novel hypothesis: that synchrony is preferred because it is easier—or, more 'fluent'—to process at a *perceptual* level.

Processing fluency is a general principle that may contribute to shaping one's preferences for a stimulus in any modality. Rather than being a feature of a stimulus, processing fluency is experienced during the act of perceiving a stimulus, and may subsequently elicit a positive affective reward (Schwarz, 2012; Winkielman et al., 2003). It may influence higher

level judgements but does not rely upon social cognition. Processing fluency judgements are general cognitive processes, and not restricted to specific perceptual domains (Reber et al., 1998, 2004; Winkielman et al., 2003). It has been suggested in a theoretical review that processing fluency may contribute towards the synchrony-bonding effect (Wheatley & Sievers, 2016). Previous studies have established that synchronised stimuli may be easier to process than non-synchronised stimuli (see Chapters 2 and 3), although whether this explains the resultant social bonding effects of synchrony remains to be empirically tested.

Synchronised stimuli may be easier to process if they are streamed or grouped together. In the visual domain, synchrony serves as a strong cue for Gestalt grouping (Rideaux et al., 2016). Synchrony is also an important cue for grouping auditory stimuli together – usually referred to as auditory streaming – as part of the process of Auditory Scene Analysis (Bregman, 1990). Identifying the source of an auditory stimulus is a challenge for perception, as sound localisation is less precise than visual localisation (Pressnitzer et al., 2011). Auditory stimuli are streamed (or grouped) together by the perceiver according to the likelihood that they originated from the same source (Bregman, 1990). The common source may be inferred on the basis of a number of cues, such as similarity in pitch or timbre, or from temporal synchronisation (Hukin & Darwin, 1995). Importantly, it is easier to process one, clear auditory stream than it is to distinguish between multiple streams (Hughes et al., 2013; Zeamer & Fox Tree, 2013), just as attending to multiple visual objects exerts greater attentional load (Alnæs et al., 2014). Synchrony may then reduce processing load as it reduces the number of stimuli to which one must attend.

Where there are multiple sources of stimuli (e.g., many auditory signals), and when an individual (the 'self') is partaking in the production of those stimuli, it may be possible that auditory signals originating from the self may be streamed together with those of others. Participants in synchrony-bonding studies often report a sensation of self-other overlap (Tarr

et al., 2014), although the cognitive mechanisms underlying this feeling remain opaque. Hypothetically, if the auditory streams originating from the self and other are similar in timing, this may lead to a perception of self-other streaming. Participants often report an experience of self-other overlap during synchrony, which could be interpreted as metaphorical or conceptual but may, in fact, be a perceptual experience. The result would be a single stream that would be easier to process than multiple streams and that would potentially be associated with relatively higher positive affect.

Previous studies by Bamford, Tarr, and Cohen (Chapter 2) aimed to establish a relationship between synchrony and processing fluency. By measuring reaction times in response to moving visual stimuli, the authors found that synchrony was easier to perceive than non-synchrony (Chapter 2). This has been extended to include participation in synchrony, by measuring reaction times in response to visual stimuli while also tapping along to an audio-visual stimulus (Chapter 3). The present study aimed to test whether reduced processing load during synchronised action is associated with higher subjective ratings of social bonding after synchrony; in essence, whether the synchrony-bonding effect may arise due to the fluency-through-synchrony effect.

Based on the principles of processing fluency and perceptual streaming, we manipulated the streaming potential of the target stimulus in a finger tapping task and measured the resultant processing load and social bonding. Streaming potential is manipulated primarily through rhythmic synchrony, but also through timbral similarity.

4.1.1 Temporal and timbral coherence

The present study aimed to manipulate fluency on a drumming task in multiple ways. Synchrony is one method of increasing fluency (see Chapter 1). The literature suggests that fluency may be influenced by many factors, such as figure-ground contrast or prior exposure

(Winkielman et al., 2003). Similar to synchrony, timbral similarity is a cue for auditory streaming (Bregman, 1990) which could reduce processing load through reducing the number of perceptual objects. Thus, the streaming potential of self and other was manipulated through both temporal and timbral coherence.

These manipulations also correspond to two types of behavioural coordination: synchrony and behaviour matching. Synchrony is a form of coordination involving the coincidence of action in time, while behaviour matching involves similarity of the kind of action being performed, without a strict temporal component or consistent periodicity (Bernieri & Rosenthal, 1991). These two types of coordination may be referred to respectively as being time-locked and kind-locked, and are not mutually exclusive (Chapter 1). Variations of both these dimensions of coordination between the participant and their virtual partner may impact upon processing fluency, potentially with an additive effect. The precise ways in which synchrony and timbral similarity are operationalised for the purposes of this study are now discussed.

4.1.1.1 Synchrony

As previous literature has noted, synchrony of timing is one of the strongest cues towards perceptual streaming, leading to multiple stimuli being perceived as originating from a single source (Hukin & Darwin, 1995). Previous behavioural synchrony studies have manipulated timing through various means such as: instructing participants to avoid synchrony (Reddish et al., 2013), giving participants different beats to move to (Bamford et al., 2016), giving unequal information to participants to enforce a leader-follower dynamic (Demos et al., 2012), and by including a confederate who intentionally moves out of time with the participant (Hove & Risen, 2009). There are also examples of virtual partners being used so that specific parameters of the virtual tapping partner's behaviour can be tightly controlled, reducing the variability that a confederate may introduce (Fairhurst et al., 2013). The use of virtual reality

(VR) expands upon this, allowing a participant to interact with 3D avatars that can be programmed to mirror the participant (Tarr et al., 2018).

Each of these studies purport to be testing the effects of synchrony vs non-synchrony, however the non-synchrony condition varies greatly between studies. For instance, Hove and Risen (2009) used a consistent 33% increase in tempo for the non-synchronised tempo condition, equivalent to a 120:160 tempo ratio, assuming the participant is tapping at 120 BPM. By contrast, Bamford and others (2016) used a 5% increase or decrease in tempo between conditions, corresponding approximately to 120:126 or 120:114. These previously defined tempo ratios appeared to be almost arbitrary, however.

Non-synchrony is often defined as having two (or more) individuals performing an action at different tempi or frequencies. Yet, even when two agents move at different tempi they will inevitably drift in and out of phase with each other, resulting in a complex, repeated rhythm where the period length is determined by the ratio between the two tempi. For example, 60 beats per minute (BPM) against 120 BPM will come back into phase every second beat, and will essentially create a 2/4 metre, while 110 BPM against 120 BPM will come into phase every eleventh beat of the 120 BPM stream.

In fact, the regularity in the coincidence of beats may be more important for bonding than similarity in tempo. Bamford et al. (2016) found that dyads of participants dancing at the same tempo, but with consistently phase shifted beats such that they never coincided, scored the lowest social ratings, which was lower than dyads dancing at slightly different tempi. A simple tempo ratio will result in more coincidences in a regular pattern, and thus a less complex temporal relationship than a complex ratio. Tempo ratio, therefore, may provide a means for

quantifying the level of complexity, and thus the processing load required, beyond a binary synchrony vs non-synchrony comparison.[16]

In the present study, participants were presented simultaneously with video stimuli of two isochronous beats and asked to tap along to one of these beats as best they could. The two beats were played by separate drummers, sitting side-by-side, each tapping their beat using their preferred hand. The 'target', with whom the participant was instructed to tap in time, always drummed at 120 BPM. This tempo was selected as 120 BPM is often the preferred tempo for many genres of music (McKinney & Moelants, 2006). The target tempo was kept constant between conditions as tapping at a faster tempo may have induced increased arousal in the participant (Maddigan et al., 2019). The 'distractor' (i.e., the non-target drummer) drummed at one of three tempi: 120 BPM, 80 BPM, or 113.5 BPM. This made three different tempo ratio conditions between the target and distractor drummers, which were defined as increasingly complex ratios: 1:1, 3:2, 256:243.[17] These tempo ratios are defined – for the purposes of this study – as synchrony, complementary, and non-synchrony. Participants experienced these three conditions as a within-subjects factor, in counterbalanced order.

[16] Rhythmic ratios have also been used for discrete rhythmic patterns of three notes as a method of quantifying complexity. Jacoby and McDermott (2017) found a consistent preference for drumming simple-ratio rhythms across two cultures (USA and Tsimané) and between musicians and non-musicians in the USA. Participants were asked to produce three-note rhythms, which were then recorded on a drum pad, in a novel method allowing participants to show their preferences through their drumming behaviour. The produced rhythms ranged from very simple patterns with three notes of equal length (1:1:1), to much more complicated rhythmic patterns with complex relationships (e.g. 2:3:3). There were also subtle cultural differences. Tsimané participants had a much stronger bias towards producing the simplest rhythms (1:1:1), while American participants had a more even distribution across the simpler rhythms (e.g., 1:1:2 to 2:3:1), however, both groups avoided more complex rhythms (Jacoby & McDermott, 2017).

[17] These specific ratios were chosen as they correspond to the frequency ratios of the Perfect Unison, Perfect Fifth, and Minor Second respectively, in the Pythagorean tuning system (Schulter & Burmeister, 2015). Previous research in pitch perception has suggested that simple ratio intervals are easier to process than complex ratios (Bidelman & Krishnan, 2009). This has not been applied to tempo ratio perception in this way, but it provides a useful starting point when trying to establish a systematic way of varying temporal coherence beyond binary contrasts between synchrony and non-synchrony.

Based on the novel processing fluency account introduced above, it is expected that simpler ratios would be more likely to be streamed together into a rhythm that would reinforce isochronous tapping. However, there is often a 'sweet spot' for complexity in aesthetic judgements; if a stimulus is too complex it is too difficult to process, but if a stimulus is too simple it becomes boring (Chmiel & Schubert, 2017). Therefore, it is also possible that a medium-complexity ratio, such as the 120:80 BPM condition used here, may be preferred over total synchrony. Ultimately, this tempo manipulation provides a means of varying the complexity of the stimuli, and thus the resultant processing fluency for the participant.

4.1.1.2 Timbral similarity

As the present study is concerned with the effects of stimulus complexity, rather than of synchrony *per se*, a manipulation of timbral similarity was also included. Timbre is regularly defined as being the quality of a sound that enables a listener to distinguish two sounds played by different instruments and which is neither pitch nor rhythm (Houtsma, 1997). At an acoustic level, an important contributing factor to the perception of timbre is the vertical alignment of frequencies that makes it a similar concept to harmony (Hasegawa, 2021). Both timbre and harmony can be used as cues for auditory streaming (Bregman, 1990), and there may be an interaction between these vertical cues and the horizontal, temporal cue of synchrony (Shamma et al., 2011).

The present study included timbral similarity as a between-subjects factor. Half of the participants were shown stimuli in which the target and distractor drummers used matching drums, while the other half were shown stimuli in which the distractor and target had a non-matching drums. Across conditions, the target always had the same drum to control the salience of the target, as prior research suggests that pitch can affect beat perception (Varlet et al., 2020). Other methods of manipulating processing fluency from previous literature were considered, such as 'figure-ground contrast', which refers to how easily the target stimulus is distinguished

from the background (Reber et al., 1998). However, this may have also made it more difficult to maintain synchrony with the target stimulus and thus may have interfered with the synchrony condition.

4.1.2 Measuring fluency and bonding

The present study is concerned with two main outcomes, and the relationship between them: cognitive load and social bonding. Both of these constructs have been measured in numerous ways throughout the literature.

4.1.2.1 Cognitive load

Very few studies have directly investigated cognitive load, or ease-of-processing, during inter-personal coordination tasks, so multiple measures were used in the present study so that they could be validated against each other. The simplest method to assess cognitive load is to ask for subjective ratings (Schmeck et al., 2015). As a self-report measure, it relies upon meta-cognitive awareness and the honesty of the participant, however it is quick and relatively straightforward to ask participants how difficult they found a task immediately after the task. Additional measures were also included in the present study, so as not to rely only upon self-report.

Continuous physiological measures, that would not interfere with the task, provide an alternative measure of cognitive load. One method is through pupillometry, as pupils dilate in response to stimuli, and the magnitude of the dilation is related to processing load (Beatty, 1982). Measuring the mean pupil size over the course of a trial has proved effective in previous studies, as a more difficult trial should have more dilation events of higher magnitude (Kiefer et al., 2016; Meghanathan et al., 2015). This has even been applied to study the technical difficulty of playing in a music ensemble (Bishop et al., 2021), although without the aim of

studying social bonding outcomes. Hence, mean pupil size was considered an appropriate measure of cognitive load for the drumming task in the present study.

More complex approaches to using pupillometry to measure processing load may be appropriate for this task, although are also more experimental. The Low/High Index of Pupillary Activity (LHIPA; Duchowski et al., 2018, 2020), uses a wavelet analysis to measure oscillations in pupil size and claims robust to many of the problems with other methods. The LHIPA provides a continuous measure, and the wavelet signature it analyses is less susceptible to interference from light levels (Duchowski et al., 2018, 2020). Although it has been validated with a range of cognitive tasks, it has not been used in a finger tapping or drumming task before, so its use here is primarily for validation purposes.

There has been much debate over the best way to measure pupil dilation (Attard-Johnson et al., 2019). A growing literature that has mapped the neurological pathways involved in controlling pupil dilation and its connection with attention control (Costa & Rudebeck, 2016; Joshi et al., 2016). It has been demonstrated that pupil size will respond to purely auditory events, indicating that these attentional processes are not limited to visual processing (Bowling et al., 2019; Zekveld et al., 2018). Some of the earliest pupillometry studies by Beatty (1982) measured task-evoked pupil dilation – the specific and time-limited response to a stimulus – however this would be inappropriate for a continuous task like rhythmic drumming as every drumbeat is another perceptual event.

One of the most common methods for assessing cognitive load is the dual task paradigm, in which performance on a secondary task indicates the cognitive resources required by the first task (Abernethy, 1988). This may be a reliable and objective measure of processing load but necessitates a more complex experimental design in which participants must complete two simultaneous tasks. Given that part of the enjoyment of synchronised movement may arise

from an induced state of flow (Stupacher, 2019; Wasley et al., 2015), a concept highly related to optimal processing fluency, a secondary task could disrupt any social-affective rewards that the present study aims to measure. In Study 2 we demonstrated that synchrony was easier to process than non-synchrony using a dual task paradigm, however we did not also attempt to measure social bonding within that same study. Study 3 thus aims to build upon the results of Study 2 by including direct measures of social bonding and including subjective ratings of difficulty, alongside pupillometry, in place of a secondary task as a measure of cognitive load.

4.1.2.2 Social bonding

Processing fluency is reported to be hedonically marked (Winkielman et al., 2003). The fluency-through-synchrony model suggests that this experience of pleasure may be associated with the target agent. The mechanisms underpinning 'liking' (pleasure gained from a stimulus) and 'wanting' (the motivation to gain or continue access to a stimulus) are well studied (Berridge & Kringelbach, 2008, 2015). There are numerous ways of measuring liking and wanting, however the most common are simple self-report questions (Pool et al., 2016). Therefore, the present study includes a simple visual analogue scale question asking how much the participant liked the tapping partner after each trial.

A variety of other measures have been used in previous studies to capture feelings of social bonding – and related concepts of pro-social behaviour, self-other overlap, social closeness etc. Many of these have been used in synchrony-bonding studies before, but one of the most common measures is the Inclusion of the Other in the Self (IOS) scale (Aron et al., 1992). This pictorial scale aims to assess how much one's concept of oneself overlaps with one's concept of the other. Although usually a discrete scale with seven levels, the variant developed for this study uses a continuous slider to be more equivalent to the other visual analogue scales. An additional question about felt 'connection' was also included.

4.1.3 Fluency-through-synchrony

The present study tested the novel theory that the synchrony-bonding effect is explained in part by processing fluency experienced during synchronised action. Participants were shown videos featuring different tempo ratios as a systematic variation of temporal coherence (synchrony [120:120 BPM], complementary [120:80 BPM], and non-synchrony [120:113.5 BPM]), and with matched or unmatched timbre as a manipulation of timbral coherence. Three questions were asked relating to social bonding (liking, connection, and IOS). A high degree of relatedness was expected between these questions and, if they were found to correlate, then they would be combined into a single measure, similar to the procedure of Tarr and others (2016). Processing load was measured using the self-reported difficulty scale and the pupillometry measures. The following predictions were tested:

- H1: Social bonding will be lower after non-synchrony compared with synchrony
- H2.1: Perceived difficulty will be higher after non-synchrony compared with synchrony
- H2.2: Pupil size will be higher in the non-synchrony condition compared with synchrony
- H2.3: LHIPA will be lower after non-synchrony compared with synchrony
- H3: Social bonding will be higher after the matched timbre condition compared with unmatched timbre condition
- H4: The effects of matched timbre and synchrony will be additive, i.e. videos in which the target and distractor play on matched timbre while in synchrony will result in higher social bonding than those with unmatched timbre in synchrony.
- H5: The social bonding effect of tempo ratio will be mediated by perceived difficulty
- H6: The social bonding effect of tempo ratio will be mediated by pupil size

Furthermore, the complementary tempo ratio was included as an exploratory condition. Being partially synchronised, it may have an intermediate effect on both social bonding and difficulty, and this will be explored in analysis. Drumming accuracy will also be measured through inter-onset intervals and potentially included as a covariate if any differences were found between conditions.

4.2 Method

4.2.1.1 Participants

A total of 82 (60 female, 22 male) participants were recruited from a university population in Australia, with a mean age of 20.04 (range = 18 to 40). Three participants were left-handed. All participants reported having normal or corrected-to-normal vision, and that they normally use a computer without needing glasses. This was checked before starting the experiment, as the eye-tracking headset is incompatible with glasses. Participants gave informed consent to take part in the study and received course credit for their time (see Appendices 6.3.1, 6.3.2 and 6.3.3). This study was approved by CUREC at the University of Oxford (Ref: SAME_C1A_19_002) and the ethics board of the University of Western Australia (Ref: 00007908).

4.2.1.2 Materials

Videos were recorded of three different people, each wearing white t-shirts against a pale-yellow background. The people used in these video stimuli were all white males, of similar age and height to reduce variability between stimuli, while still making it clear that there were different people in each video. They were only recorded from the neck down, so that they could not be identified by their face. Each of them was recorded tapping isochronous beats on a drum in time with a metronome. They were each recorded tapping at 120, 113.5 and 80 BPM. Each person was also recorded using two different drums, at each tempo. Video stimuli were recorded on a Sony RX-100 camera.

These raw videos were then combined to produce stimuli for the study. The stimulus was a split screen video consisting of a 120 BPM tapping video on the left side of the screen and a 120, 113.5 or 80 BPM video on the right (see Figure 4.1). This created three conditions based on tempo ratios between the videos: 120:120, 120:113.5, 120:80. In addition, two different timbre conditions were created for each tempo ratio, depending on whether both

subjects in the stimulus were using matching drums or not. This was done for every possible permutation of people in the videos, ensuring that every person was tapping with every other person, in every condition, but never with themselves, resulting in 36 different videos. See Appendix 6.3.11 for example stimuli.



Figure 4.1 Screenshot of a video stimulus. Participants were instructed to drum along in time with the person on the left of the screen.

The PsychoPy3 Experiment Builder (v2021.1.4) was used to construct the experiment, manage stimulus presentation, collect self-report data, and trigger the eye-tracker (Peirce et al., 2019). Video stimuli were presented on a 27" monitor with 1920x1080 resolution and a refresh rate of 60Hz, and audio was presented through standard headphones. The code for the experiment can be found in the supplementary online materials.[18]

[18] https://github.com/Social-Body-Lab/fluency-through-synchrony/tree/main/Study3/Study3_Experiment

Two different drums were used throughout the experiment: a djembe and a hand drum. Videos in the matched timbre condition had two djembe, while the unmatched condition had the target playing a djembe and the distractor playing the hand drum. Participants always played a djembe. The participants' djembe was fitted with an Adeline AD-35 contact microphone, to record their drumming. Both the audio output from the stimuli and the contact microphone were recorded by a Zoom H4n recorder.

A Pupil Core eye-tracker from Pupil Labs (Kassner et al., 2014)[19] was used to track gaze direction and pupil dilation with a single camera under the participant's right eye. Calibration was done using the default method within the Pupil Capture software and the recording was controlled through the Pupil API in PsychoPy.

Individual differences were measured using a 10-item version of the big five personality inventory (BFI-10, see Appendix 6.3.7, Rammstedt & John, 2007), and the Goldsmiths Musical Sophistication Index (Gold-MSI, see Appendix 6.3.8, Mü Llensiefen et al., 2014). General demographic questions were also included (see Appendix 6.3.9). These instruments were used for the purposes of sample description only.

Self-report questions about social bonding and perceived difficulty of the task included: 'Please rate how much you like the person on the LEFT.', 'Please rate how connected you felt to the person on the LEFT.', 'Please rate the difficulty of the task.', and an adaptation of the Inclusion of the Other in the Self (IOS) scale (Aron et al., 1992). See Appendix 6.3.5 for all task-related questions. The IOS is usually presented as a 7-point scale, however the present study used a continuous measure, in which participants were asked to 'drag the "self" circle into the position that best describes your relationship to the person on the LEFT'. This was similar to the dynamic identity fusion index, as this allows more precision than a discrete scale

[19] https://pupil-labs.com/products/core/

(Jiménez et al., 2016). All of these were continuous scales with an arbitrary range of 0 to 100. Participants could record their responses on these scales by clicking and dragging a marker to the position that best represented their response. The slider occupied 60% of the horizontal axis of the screen, positioned in the centre.

4.2.1.3 Procedure

Participants were briefed on the study and were asked to confirm that they met the participation requirements. They were also instructed to remove any items of jewellery that may have interfered with their drumming. The Pupil Core headset was then fitted by an experimenter and the calibration procedure within the Pupil Capture software was run. The experiment itself was run through PsychoPy and participants were left to work through the study at their own pace. The experimenter sat out-of-view on the opposite side of the room throughout the session.

Participants were first presented with on-screen written instructions for the task, including the opportunity to practice using sliding scales and to familiarise themselves with the drum for a short practice session with only one virtual drumming partner. See Appendix 6.3.4 for the full instructions.

Participants were then presented with six trials of the main task. Each trial consisted of a video stimulus of two drummers drumming on a drum in front of them. They were instructed to drum along in time with the person on the left-hand side of the screen. Participants were assigned to either the matched or unmatched timbre group, which determined the set of videos seen, but not the drum the participant played. Each of the six videos shown to each participant could be in any of the within-subjects conditions: synchrony, complementary, or non-synchrony. Each participant was presented with each of these conditions twice, in counter-balanced order. Each video featured a different pair of drummers, and this was randomised

between trials. The counter-balancing also ensured that each participant experienced one of each of the three conditions within the first three trials. Counter-balancing was crucial since there is often a processing fluency effect associated with repeated exposures (Winkielman et al., 2003).

Throughout each trial, the participants' pupil activity was recorded with the eye-tracking headset and their own drumming recorded by the contact microphone. After each video, they were presented with the battery of self-report, continuous-scale questions.

After completing all trials, participants completed questionnaires for personality, musical sophistication, and general demographics. At the conclusion of the study, they were fully debriefed as to the aims and hypotheses of the study (see Appendix 6.3.6).

4.2.1.4 Data Preparation and Analysis

Data were processed in Python with Pandas and Spyder. The data output from PsychoPy was imported, including the subjective measures of Inclusion of Other in the Self (IOS), Connection, Liking and Difficulty. These were all continuous variables of arbitrary units, recorded on a scale of 0 to 100. The output from the Pupil Core eye-tracker was then segmented by trial. Eye-blinks, as identified by the Pupil software, and any frame with an estimated pupil size of 0, were removed (Mathôt et al., 2018). Pupil size estimates were calculated using 3D models generated by the Pupil software. These 3D pupil size estimates were used as they are robust to the foreshortening error, meaning that no correction was required (Petersch & Dierkes, 2021). Pupil size was not baseline-corrected, as the mean pupil size was taken over the length of the whole trial (60s) consisting of many events, and baseline correction is usually applied when analysing a pupillary response to a specific event (Mathôt et al., 2018). For each trial, a Low/High Index of Pupillary Activity score was calculated (Duchowski et al., 2020), as well as a mean pupil size (Kiefer et al., 2016).

Audio data were imported. Each audio track was recorded in stereo, with one channel for the contact microphone to record the participants' tapping, and another channel to record the audio output from the computer that was heard by the participant. The audio was first segmented by identifying the beginning and end of each trial for each recording. A beat analysis of the participant channel was then conducted using Librosa (McFee et al., 2022). This was converted to inter-onset intervals (IOI), from which a mean and variance in IOI was calculated per trial. The beat analysis was used just to confirm that participants were accurately tapping along with the target drummer and that any differences between conditions could not be attributed to a difference in task performance.

Pupil recordings of less than 4000 frames were excluded, as this implied a lot of missing data within the trial: 65 trials out of 492 (13.21%) were excluded in this way. This left a total of 80 participants and 427 trials. 29 participants had missing audio data due to recording errors, however these were not excluded for any purpose other than in the analysis of IOI variance.

Correlations and an exploratory factor analysis were conducted in JASP (JASP Team, 2020), to explore the relationship between dependent variables, and to reduce the number of dimensions in the social bonding measures. Repeated-measures ANOVAs were conducted with Pingouin in Python (Vallat, 2018). These were used to check for differences between tempo ratio conditions on drumming performance (mean IOI and IOI variance), perceived task difficulty, mean pupil size, LHIPA, and social bonding. Linear models and mediation analyses were conducted in R (R Core Team, 2022). The complete dataset (including excluded trials) and the scripts for all analyses can be found in the supplementary online materials.[20]

[20] https://github.com/Social-Body-Lab/fluency-through-synchrony/tree/main/Study3/Study3_Analysis

4.3 Results

An initial correlation analysis was conducted between the subjective dependent variables and pupil size (see Table 4.1). Significant, strong, positive correlations were observed between the three social bonding measures (liking, connection, IOS). Significant, negative correlations were observed between difficulty and liking, connection, and IOS. Small but significant, positive correlations with difficulty were observed with pupil size, IOI variance, and LHIPA.

Table 4.1 *Correlations between dependent variables.*

Variable	Liking	Connection	IOS	Difficulty	Pupil Size	LHIPA
1. Liking	—					
2. Connection	0.728 ***	—				
3. IOS	0.503 ***	0.575 ***	—			
4. Difficulty	-0.302 ***	-0.275 ***	-0.102 *	—		
5. Pupil Size	0.074	0.032	0.079	0.097 *	—	
6. LHIPA	-0.045	0.004	0.067	0.126 **	-0.053	—
7. IOI Variance	-0.093	-0.071	0.017	0.131 *	-0.031	0.058

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

Given the high interrelatedness of the dependent variables, an exploratory factor analysis was conducted between the three self-report social bonding measures. These variables met assumptions of sampling adequacy and sphericity, KMO = .677, $X^2(3) = 545.529$, $p < .001$. One factor was found, explaining 62% of variance (eigenvalue = 1.866), which included high factor loadings from the three variables: connection (.912), liking (.797), and IOS (.630). Because of this, these variables were combined to create a new 'bonding' variable (Cronbach's $a = 0.798$) which was used for further analysis.

Before testing any of the hypotheses, differences on Inter-Onset Interval (IOI) variance and mean IOI were tested between conditions. These were included as control measures to ensure that tapping accuracy was not sacrificed between tempo ratio conditions. This is particularly important given the previously noted correlation between IOI variance and difficulty. Repeated-measures ANOVAs found no significant effect of tempo ratio on either mean IOI, $F = 1.874$, $p = .159$, or IOI variance, $F = 0.283$, $p = .754$, so tapping performance did not appear to differ between conditions (see Figure 4.2).

4.3.1 Effect of tempo ratio on bonding and processing load

To investigate the effects of tempo ratio on each of the dependent variables, a series of repeated-measures ANOVAs were conducted. These revealed significant effects of tempo ratio on bonding (H1), difficulty (H2.1), and pupil size (H2.2), but no effect on LHIPA (H2.3). For the full statistics and plots of these tests, see Figure 4.2 and Table 4.2. As the data for pupil size, difficulty, and LHIPA violated the assumption of sphericity, the results for these measures are reported with Greenhouse-Geisser corrections applied. Bonferroni corrected post-hoc comparisons revealed that bonding (H1) was significantly lower in the non-synchrony condition compared with synchrony, $t(78) = -6.287$, $p < .001$, and also lower in non-synchrony compared with the complementary condition, $t(78) = -5.253$, $p < .001$. Significant differences in difficulty (H2.1) were observed between all pairs of tempo ratio conditions, with non-synchrony being the highest in difficulty and synchrony being the lowest: complementary was lower than non-synchrony, $t(78) = -6.660$, $p < .001$; complementary was higher than synchrony, $t(78) = 4.761$, $p < .001$; non-synchrony was higher than synchrony, $t(78) = 8.585$, $p < .001$. In the pupil size measure, only non-synchrony was significantly higher than synchrony, $t(78) = 2.618$, $p < .05$.

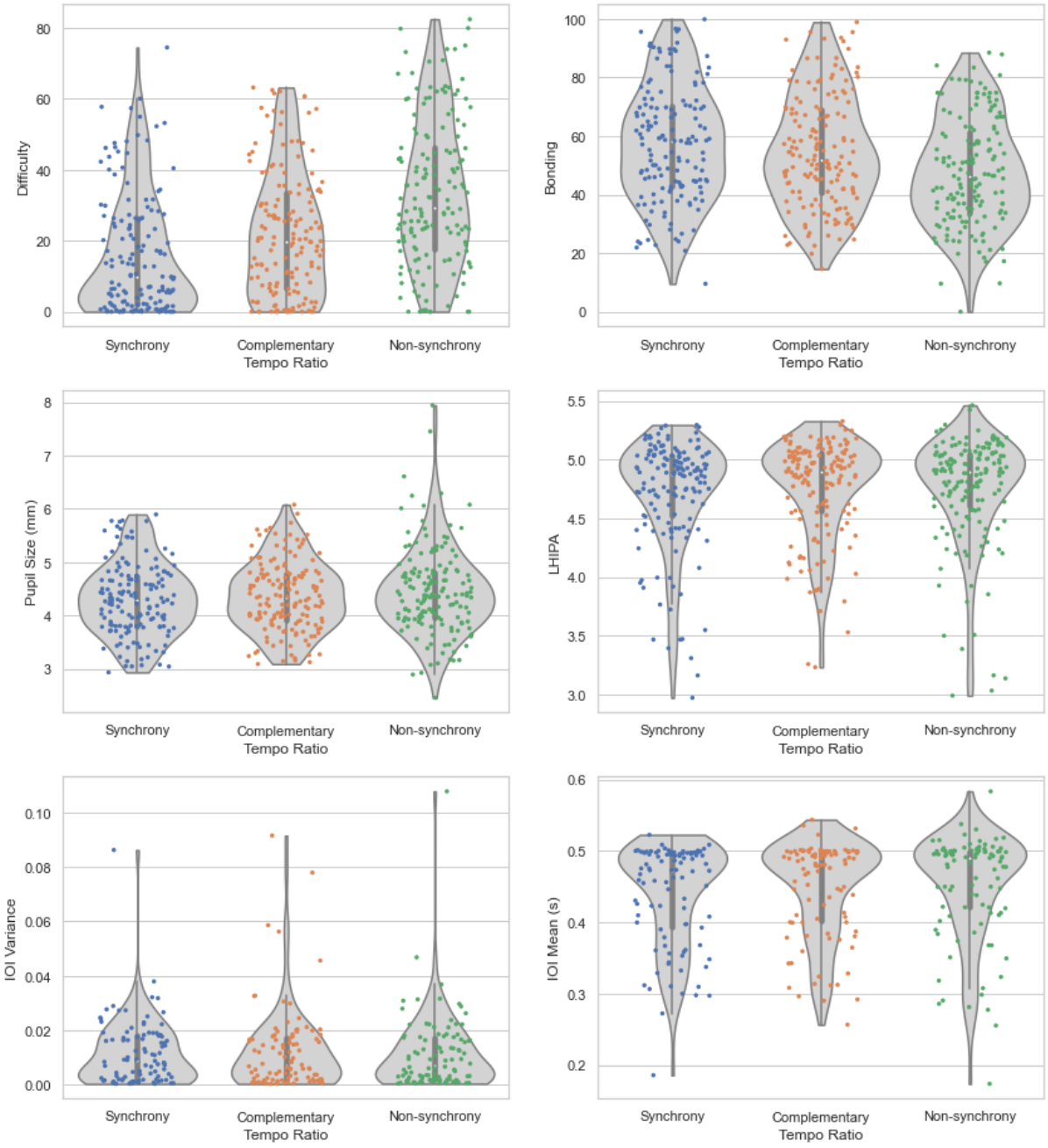


Figure 4.2 Violin plots of effect of tempo ratio on perceived difficulty, bonding, pupil size, LHIPA, IOI variance, and mean IOI.

Table 4.2 *Repeated-measures ANOVAs between levels of tempo ratio for each dependent variable.*

	df1	*df2*	*F*	*p*	η_p^2	ε
Difficulty	2	156	54.42	<.001	0.410	.791
Pupil Size	2	156	4.471	.017	0.054	.846
LHIPA	2	156	2.351	.104	0.029	.901
IOI Variance	2	98	0.2834	.737	0.005	.927
Mean IOI	2	98	1.874	.159	0.036	.992
Bonding	2	156	25.48	<.001	0.246	.959

4.3.2 Effect of timbral similarity on bonding

Data were then exported to R (R Core Team, 2022) for linear and mediation modelling. Two binary contrast variables were constructed from the three tempo ratio conditions to measure the effects of the complementary and non-synchrony conditions when compared with the synchrony condition, treating synchrony as the reference condition.

A Linear Mixed-Effects Model (LMEM) was constructed to predict bonding from the two tempo ratio comparative conditions (complementary vs synchrony, non-synchrony vs synchrony), timbre (unmatched vs matched), and their interactions as fixed factors and including participant as a random factor (Table 4.3, Figure 4.3). The only significant effect found was of the non-synchrony condition, as observed in previous analysis. Matched timbre did not lead to greater social bonding (H3) compared to unmatched timbre, nor was there any additive effect with tempo ratio (H4).

An additional LMEM was constructed for exploratory purposes, including interaction effects between the difficulty and pupil size measures, and may be found in Appendix 6.3.10.

Table 4.3 *Results of the linear mixed-effects model for the effect of tempo ratio and timbre conditions*

Predictors	Estimates	*CI*	*p*
(Intercept)	52.14	47.26 – 57.01	<.001
Tempo Ratio (Complementary)	0.27	-1.86 – 2.39	.806
Tempo Ratio (Non-synchrony)	-4.03	-6.16 – -1.89	<.001
Timbre	2.53	-4.44 – 9.50	.476
TR (Complementary) * Timbre	2.17	-0.84 – 5.19	.158
TR (Non-synchrony) * Timbre	-2.97	-5.99 – 0.04	.053
Random Effects			
σ2	137		
τ00 Participant	231.63		
ICC	0.63		
N Participant	82		
Observations	466		
Marginal R^2 / Conditional R^2	0.050 / 0.647		

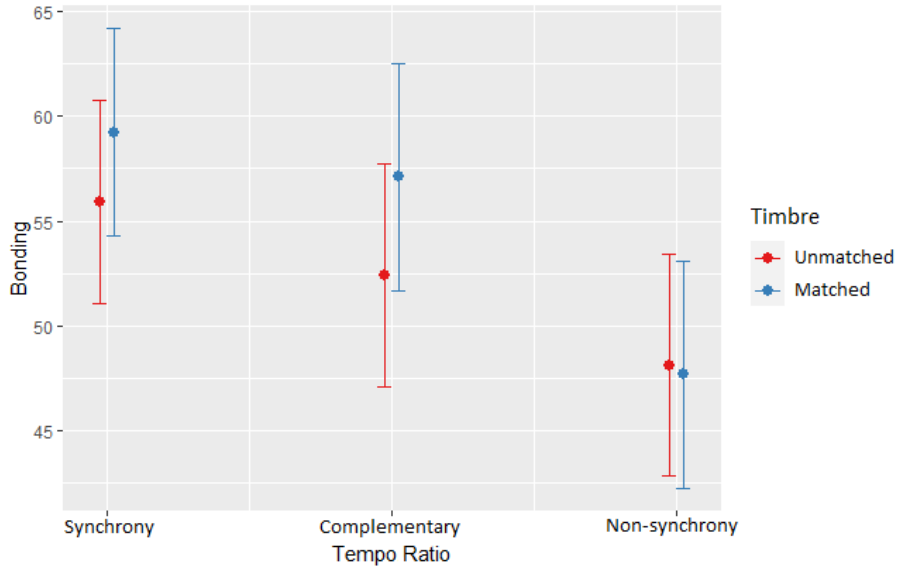
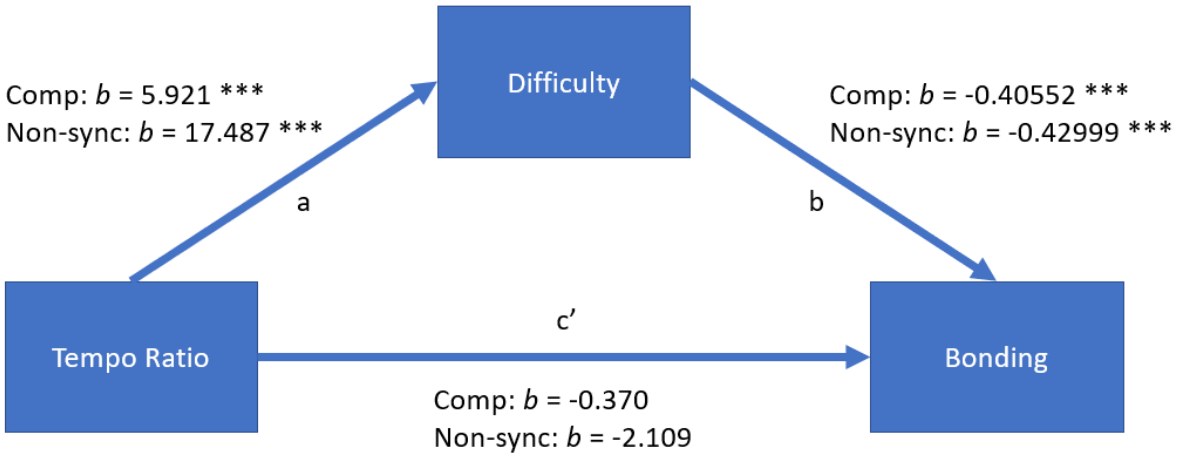


Figure 4.3 Linear mixed-effects model plotting the effects of tempo ratio and timbre conditions (same or different) on self-reported social bonding.

4.3.3 Mediation of synchrony-bonding effect by processing load

To test whether the effect of tempo ratio on bonding would be mediated by difficulty (H5), a mediation analysis was performed. As with the above LMEM, tempo ratio was divided into two binary contrast variables comparing the non-synchrony and complementary conditions to synchrony (as the reference condition). Both tempo ratio conditions were observed to significantly increase difficulty and reduce bonding. Increased difficulty also significantly decreased Bonding. The analysis found a significant indirect effect of the mediator, while the direct effect of tempo ratio was not significant, for both conditions (see Figure 4.4).

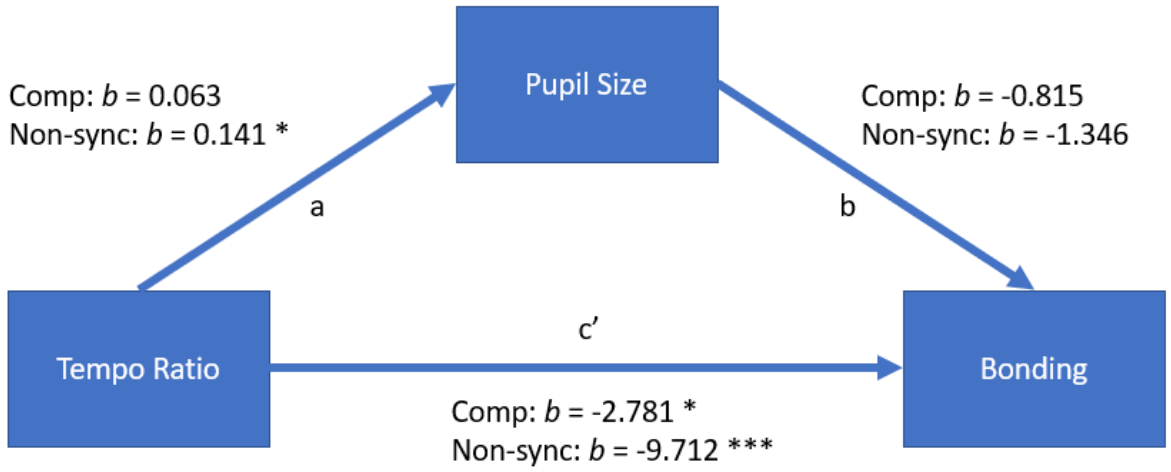


Tempo Ratio	Indirect effect (ab)	Direct effect (c')	Total effect (c' + ab)
Complementary	-2.422 ***	-0.334	-2.756 *
Non-synchrony	-7.514 ***	-2.217	-9.731 ***

Figure 4.4 Mediation model showing the direct, indirect and total effect of tempo ratio on bonding, as mediated by difficulty. Tempo ratio is split into two factors comparing complementary and non-synchrony to synchrony (the reference condition). (* $p < .05$, ** $p < .01$, *** $p < .001$)

A similar mediation analysis was conducted using pupil size as the mediating variable (H6). A significant total effect of tempo ratio was observed: bonding was lower in both the complementary and non-synchrony conditions when compared with the synchrony condition.

Non-synchrony also resulted in significantly larger pupil size when compared with synchrony. However, there was no indirect effect of pupil size on bonding, as a mediating variable (see Figure 4.5).



Tempo Ratio	Indirect effect (ab)	Direct effect (c')	Total effect (c' + ab)
Complementary	-0.045	-2.688 *	-2.733 *
Non-synchrony	-0.189	-9.555 ***	-9.744 ***

Figure 4.5 Mediation model showing the direct, indirect and total effect of tempo ratio on bonding, as mediated by pupil size. Tempo ratio is split into two factors comparing complementary and non-synchrony to synchrony (the reference condition).

4.4 Discussion

The present study aimed to test six predictions relating to the novel account of the synchrony-bonding effect through processing fluency. Most of these predictions were supported by the data and each of these are outlined below.

4.4.1 Effect of Tempo Ratio on Bonding (H1)

Adding further support to the existing literature on the synchrony-bonding effect, the present study found a significant effect of Tempo Ratio on Bonding. Participants felt closer,

more bonded, and felt more liking towards the target when the distractor was tapping in time with the target (synchrony [120:120 BPM]), compared with when the distractor was tapping out of time (non-synchrony [120:113.5 BPM]). This is consistent with previous studies that have found, on average, small, positive effects of synchrony on social bonding (Mogan et al., 2017; Vicaria & Dickens, 2016).

Although the social interaction in the present study is quite limited, previous studies have still found these effects in social bonding measures, even when participants are aware that they are working with virtual partners (Tarr et al., 2018). It should also be noted that this study specifically instructed participants to tap in time with their target, and that the interaction was unidirectional – i.e. the participant had to adapt to their target. No significant effect was found of tempo ratio on IOI variability or mean IOI, indicating that participants were similarly accurate in following the tapping of the target stimulus, regardless of tempo ratio condition. Thus, it may be assumed that participants were in-time with the target. Therefore, the difference in social bonding ratings is likely due to the distractor stimulus being in or out of time with both the target and the participant.

4.4.2 Effect of Tempo Ratio on cognitive load (H2)

The effects of tempo ratio partially supported the hypotheses regarding ease-of-processing, but there was some inconsistency between measures. Subjective ratings of difficulty were highest for the non-synchrony condition, followed by complementary, and lowest for synchrony. Consistent with the fluency-through-synchrony model, people perceived synchrony between the target and distractor to be the easiest, compared with either of the non-synchrony conditions. However, it was less clear for the pupillometry measures.

Subjective ratings of task difficulty correlated with both the LHIPA and mean pupil dilation, however the two pupillometry measures (LHIPA and pupil dilation) did not correlate

with each other. On average, pupil sizes were larger during the non-synchrony trials, compared with synchrony trials. This indicates increased processing load and more attentional resources were required during the non-synchrony trials (Beatty, 1982; Costa & Rudebeck, 2016). However, the Low/High Index of Pupillary Activity (LHIPA) found no significant effects. It is possible that some pupillometry measures may be confounded by the multiple functions of pupil dilation. Sometimes an interaction is observed between processing load and emotional arousal (Krejtz et al., 2020), so if the synchrony condition was perceived as more emotionally intense, then this could have nullified the results, although it is not clear that this should specifically affect the LHIPA and not Pupil Size.

The correlation between LHIPA and difficulty was in the opposite direction to that which would be expected: a higher LHIPA should indicate that a task required less processing, however this was not the case. The LHIPA was included as an exploratory measure. It is a relatively new measure of processing load that has been validated for a range of standard cognitive tasks (Duchowski et al., 2018, 2020), but never for a task involving rhythmic stimuli. The LHIPA relies upon a wavelet analysis of the frequency of oscillations in pupil dilation, as there is usually a higher ratio of high-frequency wavelets compared with low-frequency wavelets during tasks of higher processing load. However, this may cause issues for rhythmic stimuli. Pupils tend to dilate in response to any stimulus (Beatty, 1982) and, during the tasks in this study, participants were tapping along to a beat at 120 BPM or 2Hz. Having two perception-action events every second may have overshadowed any other pupil signals, making it difficult for the LHIPA to distinguish between conditions. This would not have been an issue for the Pupil Size measure, which analyses the average amplitude of pupil dilation, rather than timing of dilation events. It may be that the LHIPA is less suited to measuring processing load in rhythmic tasks, although further studies should confirm this.

4.4.3 Timbral similarity (H3 and H4)

As an additional between-subjects manipulation, some participants were presented with stimuli of different drums to the one the participant was playing. No significant effects of the timbre was observed, however. Our hypothesis predicted that unmatched timbre may have been more difficult to process, or less likely to be streamed together, than matched timbre. Similarity in timbre may be used as an auditory streaming cue, similar to synchrony of timing (Bregman, 1990). In theory, this means timbral similarity may also have processing fluency effects similar synchrony, and that potentially those effects may be additive. As the present study observed no effect of timbre, it may be that rhythmic synchronisation is more important than timbral similarity in both the perceived difficulty of a synchronised tapping task and in the resultant social bonding. Indeed, Shamma and others (2011) suggest that temporal coherence, or synchrony, may be the most important factor in streaming, when compared with harmony or timbre. However, it may also be that the difference in timbre between the drums was not great enough to disrupt processing fluency in any meaningful way. Future studies should investigate more possible ways to manipulate the difficulty of a rhythmic tapping task, to better distinguish between effects of synchrony *per se* and general effects of processing fluency. This could involve changes of timbre, pitch or through adding noise or distortion to the signal, for instance.

4.4.4 Bonding mediated by processing load (H5 and H6)

Finally, the present study tested the hypothesis that the effects of tempo ratio on social bonding would be mediated by perceived difficulty of the task. In this instance, the effect of perceived task difficulty appeared to mediate most of the effect of tempo ratio. This supports the fluency-through-synchrony theory, that synchronised action is easier to process than non-synchronised action, and that it is the subsequent experience of processing fluency that leads to greater liking for the other and feelings of social closeness.

An additional mediation model was constructed using pupil size as the mediating variable, in place of self-reported task difficulty. Non-synchrony was found to increase pupil size, as would be expected if non-synchrony requires increased processing load compared with synchrony. Although a weak correlation was observed between pupil size and perceived difficulty, pupil size did not exhibit the same mediating effect on social bonding. Pupil size is used here as a proxy measurement of neural activity in brain regions associated with attentional control (Costa & Rudebeck, 2016). It is possible that the perceived difficulty reported by participants, and that mediates the effect on bonding, is derived from processing load elsewhere in the brain that is not captured by pupil size. It is also possible that the effect of pupil size on bonding is too weak, given that these are very different types of measurements: both difficulty and bonding were recorded after the trial as self-report measures on continuous scales, while pupil size is a physiological measure derived from the mean pupil size over the course of a trial. Future research may be able to find better physiological or neuroimaging measurements to study the physical basis for the perceived difficulty in this task.

4.4.5 The 120-80 BPM condition

The non-synchrony condition was defined by the tempo ratio of 120:113.5, meaning that the target was tapping at 120 BPM while the distractor tapped at 113.5. An additional tempo condition referred to as ‘complementary’, with a ratio of 120:80, was included for exploratory purposes. Indeed, the results were much clearer for the non-synchrony condition, and on many measures complementary was not distinguished from synchrony. The exception was with subjective ratings of difficulty, in which the complementary condition was rated as being of intermediate difficulty between the other two conditions.

Rather than manipulating synchrony *per se*, the present study used tempo ratios to manipulate the complexity of the relationship between two auditory streams. The non-synchrony (120:113.5) condition corresponds to a ratio of 256:243, while the complementary

(120:80) condition has a simpler 3:2 relationship. The ratio of 256:243 is the frequency ratio of a Minor Second, and 3:2 is the frequency ratio of a Perfect Fifth, in Western Tonal Harmony (Schulter & Burmeister, 2015). These ratios serve as a useful starting point for quantifying the temporal coherence between two isochronous beats in terms of ratio complexity but are otherwise completely arbitrary. Using multiple tempo ratios, rather than binary synchrony vs non-synchrony conditions, may be useful for further investigating the effects of processing fluency. Tempo ratio complexity is a continuous variable which may be experimentally manipulated. The present study found that a 254:243 ratio is enough to disrupt any bonding effects of synchrony, while a 3:2 ratio has less effect. However, future research could find an optimum degree of complexity, depending on context, or that there is a certain threshold after which a tempo ratio is perceived as non-synchronous. Using ratios as a conceptual framework may prove fruitful for these future endeavours (Bowling & Purves, 2015; Larsson et al., 2019; Schwarz, 2012). This is likely an oversimplification of the real dynamics of rhythm perception. Nevertheless, the use of multiple levels of tempo ratio as a 'synchrony' manipulation is novel in the synchrony-bonding literature and has a stronger theoretical basis than a single, arbitrary non-synchrony condition.

4.5 General Discussion

The present study provides evidence for the fluency-through-synchrony model explaining the synchrony-bonding effect. The experience of synchrony in this group drumming setting was easier to process, and resulted in greater social bonding, than non-synchrony. Furthermore, the social bonding effects appear to have been mediated by the experience of processing fluency. The strength of the effect may depend on how complex, or non-synchronous, the stimuli are – defined here through tempo ratios.

Theoretically, any stimulus features that increase complexity, and thus difficulty of processing, should reduce social bonding. The present study included a variable of timbral

similarity by manipulating the drums used in the stimuli, although this had no effect. However, the difference between timbre of the drums in the unmatched condition was very mild. This was intentional, as we did not want the timbre condition to disrupt synchrony between the participant and their target, as this would provide a confound between the two experimental variables. Additionally, we decided to only use differently tuned drums in the stimuli, as the presence of other instruments could have encouraged participants to see the drummers as an in-group (Diehl, 1990). Future studies could investigate this further by using greater timbral differences or other means of manipulating stimulus complexity. Previous research into the processing fluency effect has used various manipulations of figure-ground contrast (Reber et al., 2004) which could be adapted for a replication of the present study, in which noise or distortion is added to the video.

The task used here forced the participant to be a follower, as the target drumming partner was unable to adapt to the participant. Previous research has suggested that interaction may be important to the experience of synchrony (Konvalinka et al., 2010; van der Steen & Keller, 2013), and this is certainly a feature of normal human interaction in rhythmic tasks. Future studies could allow for both the participant and target to adapt to each other, varying how easy it is to follow the target. The target could be a virtual partner programmed to adapt to the participant to varying degrees (e.g. van der Steen & Keller, 2013) and could be presented in virtual reality (Tarr et al., 2018). This would allow for a study design with only a dyad, rather than a group of three (with target and distractor) as used here. The task design used here represents a compromise between any potential variability that using human confederates as the stimuli would have introduced, while still having humans in the videos as the targets of social bonding. Following this initial study, future research should aim to replicate these results in a variety of settings.

There may be a balance to be found between complexity and simplicity. Studies of musical complexity often find an inverted-U curve, with an optimum amount of complexity (Witek et al., 2014). This may relate to a need for information gain in perception, as a stimulus which is so predictable as to be devoid of epistemic value is easily ignored (Clark, 2018). It is possible that the relationship between processing fluency and bonding is not strictly linear. Too much processing fluency, for too long, could become boring. The present study suggests that even a short exposure to the 120-113 condition impeded social bonding, however a fine-grained variation of tempo ratio would be required to find if there is a threshold for this effect. Similarly, over a prolonged period of time, the 120-120 condition may become less likable, and a slightly more complex relationship (e.g. 120-80) may become preferred. There may also be individual differences in these effects based on prior rhythmic experience from music or dance training. We collected data on prior musical training, but did not aim to recruit a wide range of participants, as individual differences were not specifically investigated in this study. Future studies could target particular groups (i.e. musicians vs. non-musicians) to investigate the effect of individual differences in musical training on the optimum level of rhythmic difficulty.

It should be noted that real music or dance settings rarely involve isochronous tapping, but have layers of complexity, and may involve synchronisation at different metrical levels, to maintain interest. In some ways, the present study resembles an ensemble setting with three drummers. The participant was always playing in-time with one of those drummers, while the third (the distractor) had a variable temporal relationship. Future studies could also examine more complex ensemble arrangements with more parts. Crucially, this line of experimentation extends the synchrony-bonding effect beyond strict synchrony and towards a conceptualisation of rhythmic coherence.

Further investigation may be required to understand why synchrony, or rhythmic coherence, might lead to processing fluency effects. Gestalt grouping and Auditory Scene Analysis (ASA) principles suggest that synchronised stimuli may be grouped or streamed together and perceived as one object (Bregman, 1990; Rideaux et al., 2016). One perceptual object provides less to process, and studies have observed that having to distinguish between ambiguous objects or streams may require greater attentional resources (Dolležal et al., 2014; Hughes et al., 2013; Zeamer & Fox Tree, 2013). This could lead to a sensation of self-other streaming, in which a person cannot distinguish between the sounds caused by their own actions and those of a synchronised partner.

Predictive processing offers an alternative level of explanation. Any actively inferring agent is testing models of the world against its perception of the world. Prediction error is then either used to update the model or is reduced through acting in the environment (Clark, 2015). In the case of two mutually interacting systems, one agent shapes the environment for the other in a perception-action feedback loop (Veissière et al., 2019). Thus, when in synchrony, one agent can continuously fulfil the predictions of the other agent. Furthermore, the predictions of an agent may be based upon the agent's own action (Clark, 2015), so the internal model of both self and other may begin to merge.

Either the self-other streaming, or model merging, described here could contribute to a reduction in processing load during actions with a high degree of synchrony or rhythmic coherence. Further study would be required to further elucidate the underlying perceptual/motor mechanisms. Crucially, the present study evades the higher-level explanations of many previous synchrony-bonding theories. Some authors have suggested a role for joint attention (Wolf et al., 2016); however, this is unlikely in the present study as participants cannot see their virtual partner's eyes. A sense of achievement of goals would be possible (Reddish et al., 2013); however, the task is not framed as a collaborative task and may

best be understood as an aesthetic experience. Furthermore, the interaction between the participant and their target is very limited, and no attempt was made to convince participants that they would ever meet the people in the videos for any further collaboration. The role of processing fluency then presents a simple, perceptual processing account of the synchrony-bonding effect which may function independently of any higher-order cognition processes associated with interpersonal synchronisation.

4.6 Acknowledgements

Thank you to the School of Psychological Science at the University of Western Australia for the use of their facilities. We also wish to thank Cal Silberstein, Andrew Silberstein, and Jake Bamford for their assistance in creating the stimuli for this study, to Arran Davis for his advice on statistics, to Sophie Berdugo and Fabienne Saurer for their helpful comments on an early draft of this paper, and to Cathy Macpherson and Bradley Walker for their assistance in the lab. The authors would like to acknowledge that the data for this study was collected on Whadjuk Ngoonar land. This project received funding from the John Fell Fund.

5 General Discussion

Synchrony is a common, and important, social behaviour. From a young age, children will spontaneously synchronise their movement to music. Once we learn to walk, we are likely to subconsciously synchronise footsteps with those around us. A wide range of the social activities that we engage with involve intentional synchrony such as music, dance, marching and rowing. Even when not trying to synchronise our movement to music, we may spontaneously do so against our will (Zelechowska et al., 2020a). The way in which we synchronise our movement to music seems also to reveal something about our music perception, as we embody metrical structures (Burger et al., 2014). Furthermore, synchronised behaviour seemingly exists in every human culture around the world (Savage et al., 2015), yet in comparatively few other species, suggesting that the ability to entrain movement to sound developed in humans since our divergence from chimpanzees (Kotz et al., 2018; Merker, 1999). One suggestion is that music and dance first emerged as simple body percussion in early hominins, relying upon an ability for motor entrainment (Montagu, 2017). Synchrony's prevalence in human societies, and relative absence in our closest relatives, raises important evolutionary questions about both the ultimate function of synchrony and the proximate mechanisms to achieve this.

One of the candidate theories is that synchrony leads to a social bonding effect, which is leveraged by music and dance behaviour (Dunbar, 2012b; Savage et al., 2021). Most primates are highly sociable, which comes with many benefits as groups that can cooperate well are often more successful than a lone individual or groups that cooperate less successfully (Okasha, 2006; Wilson, 2012). However, intra-specific cooperation also poses many cognitive challenges which may have driven the evolution of human cognitive abilities and social behaviour (Dunbar, 2009). One of the biggest challenges in large social groups is knowing who to trust (Nowak, 2006). In most primates, social grooming is the primary behaviour for

establishing social bonds and trust within a group (Dunbar, 1991). However, grooming is time-inefficient and so the suggestion is that humans largely replaced social grooming with synchronised behaviours such as music and dance (Dunbar, 2012a, 2012b). Crucially, one can sing or dance with many more people simultaneously than one could groom, making these a more efficient way of bonding with many people at once (Launay et al., 2016; Robertson et al., 2017; Weinstein et al., 2016). This effectively frames interpersonal synchrony as grooming at a distance. A range of activities that enable synchrony – including music, dance, some rituals and many sports – may have then evolved around this effect.

Music provides a sensory scaffolding to reinforce coordination between people. The synchrony that arises is then considered the 'active ingredient' behind the social bonding effects of group music-making. This synchrony-bonding effect has been demonstrated in simple finger tapping experiments (Hove & Risen, 2009), as well as more naturalistic studies in which synchrony is manipulated between dancers (Bamford et al., 2016; Tarr et al., 2016; Woolhouse et al., 2016). Meta-analyses of synchrony-bonding studies have subsequently found a small but consistent effect of synchrony on social bonding measures (Mogan et al., 2017; Rennung & Göritz, 2016; Vicaria & Dickens, 2016). It has been suggested that synchrony leads to a sensation of self-other overlap and that it engages reward systems in the brain (Tarr et al., 2014). However, exactly why synchrony should be experienced as rewarding has remained unclear.

5.1 Synchrony-bonding theories

Numerous theories have attempted to explain the mechanisms behind the social bonding effects of synchrony. Based upon ethnographic observations, Durkheim (1995) first suggested that collective action in group ritual settings may lead to an experience of *collective effervescence*: a shared, positive, emotional state. Experimental evidence for this has been mixed. Wiltermuth and Heath (2009) found that shared, positive emotions were not necessarily

required for synchronised action to increase cooperation. This finding has since been replicated in multiple studies, with one meta-analysis finding a social bonding effect of synchrony, independent of positive affect (Mogan et al., 2017).

Another theory is that synchronised action is a display of cultural affiliation or otherwise is rewarding through participation in shared culture (Clayton, 2016). Indeed, some research has suggested that some shared cultural or religious understanding of the coordination task may be required for subsequent bonding effects (Cohen et al., 2014). However, other studies have still found an effect of synchrony without a religious context, so this cannot be the only explanation (Mogan et al., 2017). Nevertheless, this does highlight the need to consider the context and intentions behind synchronised behaviour.

Mention must be made of homophily, which is a general theory that people tend to like others who are like themselves (McPherson et al., 2001). There is some evidence to suggest that people perceive someone with whom they are in synchrony as being similar to themselves (Valdesolo & DeSteno, 2011), and people often perceive a dyad as being more coordinated if that dyad is similar in other ways (Macpherson, Fay, et al., 2020). However, others have failed to find an effect of synchrony on perceived similarity (Cohen et al., 2014), so the evidence is mixed.

Some theorists have suggested that the achievement of shared goals may be crucial to the synchrony-bonding effect (Cross et al., 2021; Reddish et al., 2013). The reasoning is that if two individuals can successfully achieve coordinated action in one domain, they may be more likely to trust each other to cooperate on other tasks. Indeed, previous studies have found that participants are more likely to cooperate with (Reddish et al., 2013), and imitate (Cross et al., 2021), each other after synchrony, but only when it is goal directed. However, when Launay, Dean and Bailes (Launay et al., 2013) included a condition in which participants had the goal

of avoiding synchrony, there were no subsequent bonding effects. This implies that it is not enough to share any goal, but that there is an effect of synchrony in particular.

It is possible that some attention processes are involved. Synchronised actions tend to attract greater attention than non-synchronised actions (Large & Jones, 1999). This has been demonstrated in a wide variety of settings: in dance (Bamford, 2017), while hand waving (Macrae et al., 2008), and in 14-month-old infants are bounced in time with an adult (Cirelli et al., 2014). Extended periods of synchronised action may be an honest signal that one participant is attending to the other (Bispham, 2018). Furthermore, this maintained attention may improve one's memory for the other (Miles et al., 2009; Woolhouse et al., 2016). Whether through improved memory or signalling, synchrony may capture attention, however it is not clear why this should lead to increased social bonding.

Extending upon the existing cognitive theories is the suggestion that synchrony induces an experience of self-other overlap. Maintaining synchrony requires the maintenance of internal representations of the other's actions (Knoblich & Sebanz, 2006), which may engage perception-action networks involved in motor mirroring (Overy & Molnar-Szakacs, 2009). This experience is often likened to the rubber-hand illusion, in which a rubber hand is incorporated into one's concept of self through visual-tactile sensory synchrony (Botvinick & Cohen, 1998), although in the self-other merging account this occurs with another person (Tarr et al., 2014). The subjective experience of self-other overlap is often measured using the Inclusion-of-Other-in-the-Self (IOS) scale (Aron et al., 1992). However, self-other overlap often appears more as metaphor than mechanism, and the process through which it would lead to increased liking and trust of another is imprecise. It is possible that this explanatory gap may be bridged by incorporating the general principle of processing fluency.

5.2 Fluency through synchrony

This thesis develops and tests the explanation that synchrony may be easier to process than non-synchrony. The phenomenon of processing fluency has been observed for a wide range of stimuli, and its principles have been applied in many fields, from aesthetics to advertising (Reber et al., 1998; Reber & Unkelbach, 2010). The processing fluency principle suggests that the mind is inherently driven to reduce energy costs, and so is rewarded for perceiving an 'easy' stimulus. There are many factors that may contribute to a stimulus being perceived as 'easy'. Some of these arise from the perceiver, such as prior experience with the stimulus or others like it, while other factors may be contributed by features of the stimulus itself – e.g., simplicity of form, or symmetry (Reber et al., 2004). In a somewhat circular heuristic, the rewarding feelings that arise from perceiving an easy-to-process stimulus, may then be used as a cue for familiarity – i.e., stimuli that are familiar are easy to process, therefore the experience of processing fluency indicates that this stimulus is familiar (Schwarz, 2012). Rather than relying upon memory for a stimulus, the processing fluency heuristic allows one to make judgements about the familiarity of a stimulus. However, when processing fluency arises from stimulus features, rather than from prior exposure to a stimulus, this can lead to an illusion of familiarity. It has been suggested that such a processing fluency heuristic contributes to the affective rewards and feelings of familiarity experienced in the synchrony-bonding effect (Wheatley & Sievers, 2016). When someone synchronises with another they may create the illusion that they are familiar to the other, by making themselves easy to process. Testing processing fluency during synchronised tasks poses significant methodological challenges, as most measures of load use reaction times, which may be confounded by rhythmic tapping. To solve this we needed a physiological measure of load that could be used during a tapping task.

Throughout this thesis, I have tested a novel processing fluency account of the synchrony-bonding effect. The Study 1 (Chapter 2) used an animated flanker task and found

that participants were faster to correctly identify the direction of the target stimulus when the flanker stimuli were also moving at the same rate as the target. This indicated that synchrony is easier to perceive than non-synchrony, at least in this highly controlled setting.

Study 2 (Chapter 3) aimed to test whether participation in synchronised action would be easier to process than non-synchronised action. This was done using a dual-task experiment, in which participants tapped along to a video stimulus while completing a visual search task. We observed that participants had worse performance on the visual search task during non-synchrony. In addition, participants exhibited greater pupil-dilation during non-synchronous trials. Taken together, these first two experiments suggest that synchrony may indeed be easier to perceive and participate in than non-synchrony.

Finally, Study 3 (Chapter 4) aimed to study whether the ease-of-processing experienced during synchrony contributes to subjective feelings of reward and bonding. This was a partial replication of Study 2, however without the visual search task and with subjective ratings after each trial. In this study, synchrony was operationalised through tempo ratios, with 120:120 BPM being considered synchronous, 120:80 BPM as complementary, and 120:113.5 BPM as non-synchronous. A significant effect was found of synchrony on perceived difficulty of the task, pupil dilation, and the social bonding measures. Furthermore, perceived difficulty entirely mediated the effect of synchrony on social bonding.

These three studies begin to establish a new theoretical framework for the synchrony-bonding effect, based upon the general perceptual principle of processing fluency, while also establishing the usefulness of pupillometry in research on rhythmic coordination. For an individual, synchronising involves matching the actions of another through prediction and adaptation (Keller et al., 2014). However, when two or more agents synchronise with each other, they create a perceptual experience of congruency between their action and perception.

The act of synchronising may be reframed as an act of making oneself easier to process, thereby eliciting a reward response in those we synchronise with.

Indeed, some authors have begun to frame synchrony as a process of making oneself more predictable (Koban et al., 2019; Konvalinka et al., 2010). By synchronising with others, one may then capitalise upon a pre-existing drive to reduce processing load (Reber et al., 2004). The resulting behavioural alignment may also promote cognitive/emotional alignment (Shamay-Tsoory et al., 2019), as those who synchronise together may actually come to share cognitive/emotional states through the process of sharing behaviour. Importantly, this process of coordination is not limited to music and dance, but may be a domain-general, embodied aesthetic process (Vuoskoski & Reynolds, 2019). Processing fluency, as a general principle from neuro-aesthetics (Reber et al., 2004), provides a new framework for understanding the perception of synchrony and its resultant social effects. Crucially, it suggests the social bonding effects from synchrony may in fact arise from synchrony-induced processing fluency, rather than synchrony *per se*, meaning that other factors which could increase fluency may also increase bonding during a coordinated activity. From this basis, it is worth assessing how processing fluency integrates with other theoretical mechanisms behind the synchrony-bonding effect.

5.3 Integration between theories

Rather than relying upon any synchrony-specific mechanism, the fluency through synchrony account is based on general perceptual mechanisms. It proposes that people leverage these general mechanisms when they synchronise with others, in order to manipulate the other's perception to make the self more likable. Although this is a theory grounded in perceptual processing, other stages of cognition may also be involved, potentially creating feedback loops, and invoking some of the other theories that have previously been proposed as synchrony-bonding mechanisms.

Processing fluency may interact with the attention and motivation theories discussed earlier. Previous research has suggested a role for social attention in the synchrony-bonding effect (Wolf et al., 2016). Generally, stimuli which are easier to process attract less attention than novel stimuli (Johnston et al., 1990). However, experiencing unexpected fluency for a stimulus also results in orienting attention towards that stimulus (Reber et al., 2004; Schwarz, 2004). Other humans are usually a complex stimulus that require a lot of cognitive resources to perceive and interact with (Dunbar, 2009). Another human moving in synchrony with oneself would then make for an unexpectedly easy stimulus, which would grab the perceiver's attention and form the basis for ongoing interaction. Synchrony may also lead to increased perception of similarity, although the evidence for this is mixed (Vicaria & Dickens, 2016). Perception of similarity may also lead to increased social attention (Kawakami et al., 2021). Therefore, it's possible that processing fluency may be interacting with attention processes, either directly or mediated by perceived similarity. Achieving synchrony could guide attention towards the other, and this increased attention then motivates further synchrony in a positive feedback loop.

An initial, positive interaction through coordinated action may then form a basis to continue the behaviour. Miles and others (2010) found that a negative first impression of someone can discourage later entrainment with them. An initial experience of synchrony with a person may, however, create a positive impression, therefore generating motivation for ongoing coordination. If all members of a group experience this, then it may create the shared intentions that Reddish and others (2013) discuss. People usually seek out experiences that reduce their own cognitive load (Reber et al., 2004),[21] and people who discover others who can do that for them may be mutually motivated to seek out further interaction with those people.

[21] Although, see section 5.6.2 for more discussion of the optimum levels of cognitive load. It may be possible for a stimulus to be too easy to process.

In this way, synchrony may lead to an achievement of a shared goal (the goal being to synchronise). Achieving shared goals may then encourage future cooperation, including continued coordinated action.

This usually occurs within a social/cultural context, which may add to the processing fluency effects of synchrony. Coordinated actions often occur as part of a cultural event and achieving synchrony may be considered an integral part of participation (Clayton, 2016). Furthermore, it is possible that some activities may create an expectation of synchrony or an explicit assumption that synchrony indicates strength of bonding (Atwood et al., 2020). Music and may be used in many contexts, but particularly to encourage group synchrony through dance (Mehr et al., 2019; Savage et al., 2015). An initial experience of processing fluency due to synchronised action may have led to the creation of cultural frameworks that support further opportunities for continued synchrony. This may further drive expectations that synchrony leads to bonding, creating a positive feedback loop that amplifies the effect. Alternatively, these cultural factors could sit beside processing fluency as an additional effect. Either way, there would be an expected interaction between cultural expectations as a higher-order cognitive factor, and processing fluency for the perceptual experience of synchrony.

Finally, processing fluency is ‘hedonically marked’ and is usually accompanied by positive affective rewards (Winkielman et al., 2003). Although the research on positive affect arising from synchrony is mixed, it has been suggested that affect may become a more important factor with increasing group size (Mogan et al., 2017). If many people are mutually reducing processing load for each other, this may indeed induce a positive emotional state for those involved. This could lead to an experience of collective effervescence that some have described (Durkheim, 1995). The positive affective rewards may then be associated with others participating in the interaction. However, further research may be required to fully understand

all the mechanisms involved in the synchrony-bonding effect as it operates in different contexts.

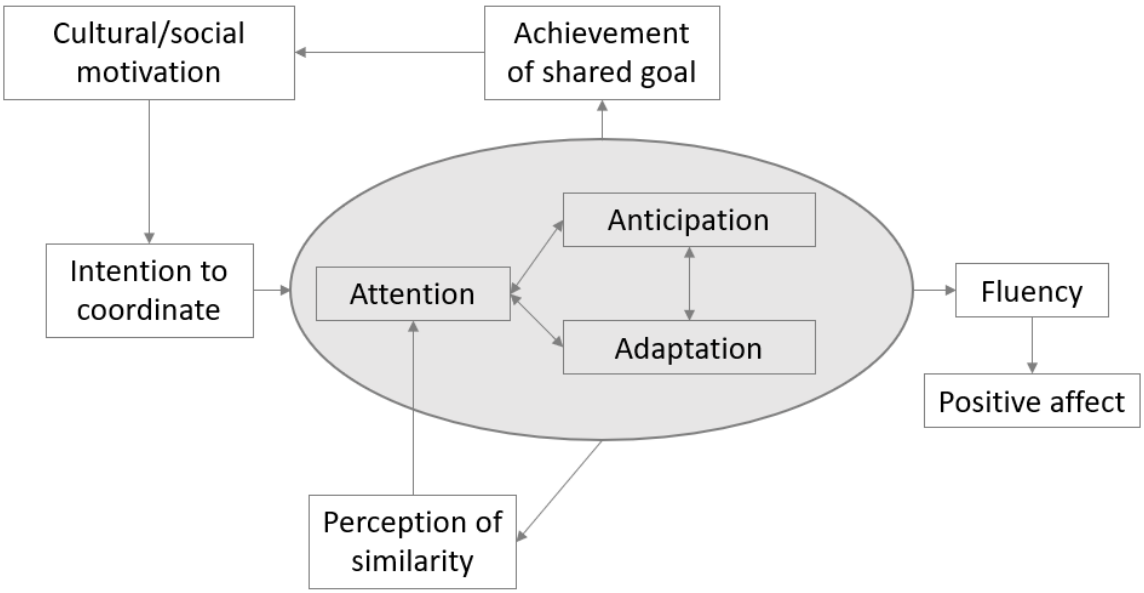


Figure 5.1 Proposed model outlining the possible relationship between different synchrony-bonding theories.

Based upon all the various mechanisms described above, a theoretical model could be constructed to illustrate how these theories may interact. In Figure 5.1, the three fundamental cognitive processes required for maintaining synchrony are placed in the centre, as suggested by Keller (2013). When synchrony is achieved, that will increase the perception of self-other similarity, which further guides attention towards the other, making it easier to synchronise. Depending upon the intention of the participants, achieving synchrony may also represent an achievement of a shared goal, which is rewarding and may motivate further collaboration. This could be additionally driven by social/cultural factors depending upon context and upon any expectations that participants may have about synchrony. As interaction continues and synchrony is maintained, processing load for participants should be reduced, leading to an

experience of processing fluency which creates positive, affective rewards that are associated with the other. Finally, it is possible that through such coordination the participants are able to make more accurate predictions about each other's cognitive/emotional state, leading to increased empathy, a possibility which is explored more in section 5.6.3.1. This is largely speculation, as the empirical work to establish such a model remains to be done. Nevertheless, this model does provide a potential pathway through which processing fluency may underpin, and be amplified by, many of the other mechanisms proposed in the synchrony-bonding literature.

5.4 Deeper mechanisms

The studies contained in this thesis demonstrate that synchrony is easier to process than non-synchrony under given circumstances. As previously discussed, this fluency-through-synchrony effect may be small, it could set into motion feedback loops that have larger implications. However, the studies presented here do not elaborate on exactly why synchrony reduces processing load. To do so, further studies into deeper mechanisms would be required.

Maintaining synchrony through entrainment to an external beat is a complex cognitive task involving a range of processes. Keller (2013) proposes three main cognitive processes during inter-personal coordination tasks: attention, anticipation and adaptation. In Chapter 1 we discussed how synchrony may reduce the cognitive load on each of these processes. Although the empirical studies presented in this thesis do not directly address the distinction between these processes, it is worth returning to consider them in light of our findings.

5.4.1 Attention and streaming

Synchrony is a reliable cue for grouping objects together in both auditory (Bregman & Campbell, 1971; Shamma et al., 2011) and visual (Rideaux et al., 2016) domains. Although in Studies 1 and 2, participants were not asked about the number of perceptual streams they were

perceiving, those studies were previously discussed as reducing load through a reduction in the number of objects needing to be processed. In particular, the fact that there was very little difference observed between the single hand and synchrony conditions in Study 2 is consistent with the idea that the synchronous hands were perceived as one object. Tracking multiple objects has been associated with increased processing load (Alnæs et al., 2014), so being able to group objects together may effectively reduce processing load.

This perceptual experience may also serve to elaborate on the self-reported experience of self-other merging. Participants often report a sense of self-other overlap following synchronous activity (Tarr et al., 2014), although it is often understood as a metaphor. However, if two members of a dyad are drumming in synchrony with their drumbeats in tight temporal alignment, then their auditory perception may process both their own drumbeats and those of their partner in a single auditory stream, according to Auditory Scene Analysis (Bregman, 1990). This could lead to an experience of self-other streaming, thus reframing self-other merging as a perceptual experience as much as a social one. Viewing self-other overlap in this way would support comparisons with the rubber hand illusion, which is similarly based in perceptual coincidence (Botvinick & Cohen, 1998). This may further relate to the concept of entitativity, in which those who move in synchrony are perceived as one entity (Lakens, 2010). Given the perceptual grouping effects of synchrony, this may well be the case.

5.4.2 Anticipation and prediction

In addition to reducing attentional load, synchrony is also predictable. This is especially true for the isochronous tasks used in this thesis, and regularly in the synchrony-bonding literature, although most rhythms have a predictable temporal structure which may make them aesthetically appealing (Vuoskoski & Reynolds, 2019). Predictive processing has been used to explain a wide range of behaviours since the formalisation of the Free Energy Principle (Friston & Kiebel, 2009). It conceptualises cognition as a continual process of making

predictions about the environment and then testing those predictions against perception. In active inference, an agent has two option when prediction errors are perceived: either update the internal model with the new information, or take action in order to change the environment to match the internal model (Friston & Frith, 2015b). For example, if an agent expects a beat at a particular moment within a repetitive rhythmic pattern and that beat does not occur, a prediction error will arise. The agent then either updates their model of the rhythm so that they will not expect a beat there in the future, or they produce the beat they were expecting themselves to fulfill their own prediction. This theoretical framework is potentially useful in the context of inter-personal coordination, and fits well within the observed processing fluency effects of synchrony.

Active inference is fundamentally embodied, as the agent interacts with its environment, through perception and action, through its body. Predictions are often made in relation to the body (Clark, 2015). Intra-specific communication of any kind involves the interaction between two or more mutually inferring agents with similar bodies (Friston & Frith, 2015a); they make predictions about each other, and provide the perception for each other. Synchronised interaction then provides a special case in which perception of the other matches one's own action. Thus the other's actions can be easily predicted from one's own actions.

This could provide the basis for the observed processing fluency effects of synchrony. Just as processing fluency is rewarding (Winkielman et al., 2003), an active inference agent may experience affective rewards from prediction fulfilment (Hesp et al., 2019). It has been observed that synchrony may emerge spontaneously from mutually inferring agents (Koban et al., 2019; Palacios et al., 2019). Two agents that wish to minimise prediction errors for both themselves and their conspecifics are thus likely to fall into synchrony, generating rewards for all involved.

Active inference is a newer theory than processing fluency, but both highlight a drive for agents to reduce processing load. Active inference primarily differs from processing fluency in that it is derived from the mathematical formalism of the Free Energy Principle (Friston & Frith, 2015b). While processing fluency operates as a general principle stating that agents will seek to reduce processing load, active inference can use FEP modelling to predict how an agent will act to reduce free energy. It is also not incompatible with the streaming model presented above, as perceptual streams that are more predictable relative to each other are more likely to be streamed together (Bendixen, 2014). Active inference, therefore, provides a more precise model explaining why synchrony may be easier to process: synchronous stimuli are more likely to be streamed together and perceived as easier because they are more predictable.

To truly test an active inference account of the synchrony-bonding effect, a generative model would need to be tested against empirical data. That was beyond the scope of the studies contained within this thesis, which only aimed to establish whether processing load plays a role in the synchrony-bonding effect. Further research may explore whether this reduction in processing load is due specifically to the increased predictability of synchrony relative to non-synchrony. Nevertheless, the generality of the active inference account is provocative, and would have wider implications if it can be applied successfully.

Considering synchrony as an active inference task has some wider implications. It recontextualises interpersonal interaction as an opportunity for information gathering (Clark, 2018). Within active inference, agents have hidden states that are inaccessible to other agents. However, inferences may be made about another agent through taking action and observing their reaction. Such interaction has been termed ‘epistemic foraging’ (Clark, 2018). Music and dance then provide ideal grounds for epistemic foraging, as agents mutually agree upon a framework for the interaction (i.e., musical and rhythmic structures). Shared knowledge of the task, combined with shared intentions, then serve as shared priors on which to base the

interaction (Vasil et al., 2020a). The development of these shared knowledge structures to enable inter-personal interaction could even serve as the basis for the development of culture (Veissière et al., 2019). Thus the cultural evolution of music could be explored through the lens of such epistemic foraging between groups of mutually inferring agents.

Another possibility is to view the experience of self-other overlap through the lens of active inference. It has been suggested that a concept of self develops through experience of taking actions and testing predictions about the outcome of those actions (Kahl & Kopp, 2018). One's own actions become more predictable over time, and so one can make more temporally precise predictions about oneself (Kahl & Kopp, 2018). It is possible that predictability itself may be used as a cue to self-identification (Apps & Tsakiris, 2014); stimuli that result from one's own actions are usually highly predictable, therefore a stimulus that is very easy to predict is more likely to have resulted from one's own actions than a less predictable stimulus. This is also seen in neurophysiological recordings, as neural responses to the perceptual consequences of one's own actions are usually attenuated (Brown et al., 2013; Gao & Gozli, 2021). Crucially, the response to one's own actions is also attenuated during interaction with other agents (Weiss et al., 2011), possibly an expression of the prioritised integrative attention noted by Keller (2001) in musical ensemble performance. If the other's actions match one's own, however, then they may be attenuated in a similar way, and one may infer that those actions originated with the self. Given that the predictability of one auditory stream relative to another can be used as a cue for streaming them together (Bendixen, 2014), the increased predictability of interpersonal synchrony may contribute to the experience of self-other streaming.

5.5 Constraints and Limitations

There were several limitations with the empirical studies in this thesis. As the first body of work attempting to systematically measure processing load during synchronised behaviour, there were few experimental paradigms to draw upon. Processing load is often measured using

reaction times (Reber et al., 2004) or by measuring the effect of a primary task on a secondary task (Abernethy, 1988). Reaction time measures are difficult to implement in rhythmic tasks, as participants will often try to respond in time with the beat (Killingly et al., 2021). This issue was avoided in Study 1 through using non-rhythmic stimuli, but this means the results from Study 1 are more difficult to generalise to rhythmic synchronisation, as discussed in Chapter 2.

Study 2 measured processing load through reaction times on a secondary task. The task had been chosen so that reaction times should be longer than a single beat, to prevent participants from responding on every beat. This method proved fruitful, however it still has limited ecological validity, as normal synchronised action does not involve a visual attention task overlayed upon one's synchronisation partner. There is also scope to replicate this with different secondary tasks, as Study 2 only used a visual search task, and there may be other cognitive processes involved other than just visual attention. Different secondary tasks could be used to determine exactly which cognitive processes are most impacted by non-synchrony.

Eye-tracking was used as an exploratory measure in Studies 2 and 3, although this proved less conclusive. This is likely because every beat would elicit a pupil dilation event (Bishop et al., 2021) which may have interfered with the LHIPA as a wavelet analysis. Mean pupil size did find significant results, although this is a rather crude measure. Other physiological measures of processing load that could be included during a synchronisation task should be explored for future studies.

There is also scope for using different conditions in any future replications of Studies 2 and 3. Study 2 included three conditions: synchrony, non-synchrony, and control. The control condition, in this instance, had only one hand moving. This was for the purpose of investigating a possible streaming mechanism, as if there were no difference between two hands tapping in

synchrony and only a single hand, then it may be that the two hands in synchrony were streamed together. However, it would have also been useful to include a condition with no hands tapping, to establish a baseline performance level on the visual search task alone. Meanwhile, Study 3 included a manipulation of timbral similarity which found no effects. This is likely because the difference in timbre was not great enough to be disruptive in any way. Timbre was included as an exploratory variable, but future research could use multiple levels of timbral difference to investigate whether a greater difference could have an effect.

It is important to consider the studies within this thesis within the context of the broader synchrony-bonding literature. The findings here are broadly consistent with meta-analyses that have found a small, positive effect of synchrony on social bonding outcomes (Mogan et al., 2017; Rennung & Göritz, 2016; Vicaria & Dickens, 2016). Some of these analyses have also investigated the role of mood and emotion, finding that synchrony has very limited effects on mood and emotion (Mogan et al., 2017). This is something that should be explored further, as processing fluency often does have short-term, positive effects on affect, by engaging reward systems in the brain (Winkielman et al., 2003), however these effects may be too short-term to be captured in the self-report emotion scales used in previous synchrony literature. Vicaria and Dickens (2016) also suggest that there may be some effects of gender, and of how well participants know each other before coordinating in a dyad. Given that our studies always used video stimuli of a person that the participants had never met, there is scope for future research to investigate how prior knowledge may influence processing fluency effects arising from synchrony. Furthermore, our video stimuli only included male drummers, while most of the participants were female. Some studies have suggested that the greatest synchrony-bonding effects occur in female-female dyads (Vicaria & Dickens, 2016). Gender of the stimuli was kept controlled in our studies to reduce the number of factors, but future research should investigate this explicitly.

One issue for the synchrony-bonding literature in general is the possibility of experimenter bias and expectancy effects. Rennung and Göritz (2016) found that a significant number of synchrony-bonding studies failed to control for this, which may have artificially magnified the observed effects. This is not a concern for the studies in this thesis, as conditions were always counter-balanced through the software and the experimenter was always blind to the condition that a participant was in at any moment. Furthermore, it has been suggested that the prosocial effects of synchrony may be largely attributed to expectancy effects (Atwood et al., 2020). However, there are numerous issues with the findings of Atwood and others (2020). Their study is a partial replication of Wiltermuth and Heath (2009), which used highly social tasks embedded in a salient cultural framework. By contrast, the studies in this thesis used very abstract tasks, removed from most social context. Furthermore, these studies were mostly measuring processing load, and mostly as an implicit measure (with the exception of Study 3 which asked about self-reported task difficulty), and we see no reason why participants would expect synchrony to reduce processing load. Participants were certainly blind to this hypothesis. While the limited ecological validity of the present studies is a limitation on one hand, on the other hand it mitigates the likelihood of expectancy effects. Atwood and others (2020) highlight the importance of running high quality, double-blinded experiments. Some studies in the synchrony-bonding literature have not maintained this standard, however this does not invalidate the area of research as a whole.

5.6 Further implications and future research

So far, I have outlined the evidence for the fluency through synchrony account of the synchrony-bonding effect, how it may integrate with some of the pre-existing theories about synchrony-bonding mechanisms, and some further possibilities for conceptualising the processing load of synchrony. For the remainder of this chapter, I will discuss the broader

contributions of the fluency through synchrony account, and some future directions for this research.

5.6.1 Individual differences

The wider context in which synchrony takes place is starting to be investigated. Some social-motivational factors contribute to whether people may wish to synchronise, as has already been discussed (Miles, Griffiths, et al., 2010). Certain personality traits may also correlate with the way people embody rhythms (Agrawal et al., 2022; Bamford & Davidson, 2019; Carlson et al., 2016), as can emotional state (Van Dyck et al., 2014), and mental health symptoms may influence how one relates to others in a coordinated task (Macpherson, Marie, et al., 2020). However, it remains unknown how these factors may mediate the experience of processing fluency arising from synchrony, which would be a question for future studies.

5.6.2 Rhythm ratios and optimum complexity

Most of the existing work on processing fluency presumes that less processing load is always the most rewarding. However, inverted-U curves for aesthetic preferences are very common (Chmiel & Schubert, 2017). Some theorists would also suggest that any stimulus which is too predictable (as in perfect synchrony) will inevitably lose interest, and that some degree of surprise is required to maintain attention (Clark, 2018). Indeed, any social interaction may be framed as a process of learning about the other through balancing alignment (e.g., behavioural synchrony) and surprise (Vasil et al., 2020b; Veissière et al., 2019). This has led some authors to begin searching for a balance between synchrony and complexity (Ravreby et al., 2022; Richmond, 2019). This is an exciting new direction for synchrony-bonding research.

Most of the existing studies on the synchrony-bonding effect have simply compared synchrony to non-synchrony (Mogan et al., 2017). Non-synchrony is often defined by different tempo relationships between stimuli, although there has been little systematic variation of this.

Study 3 used three different tempo ratios (120:120, 120:80, 120:113.5). The next steps would then be to find an optimal level of processing load, which may in fact be dynamic depending on context. The use of tempo ratios may prove a useful measure of complexity for isochronous tapping tasks, although others have suggested some utility in rhythm ratios for simple rhythmic patterns (Jacoby & McDermott, 2017). More sophisticated metrics would be required for more naturalistic synchrony tasks. Future studies could involve synchronous tapping tasks with varying tempo ratios to find an optimum, as measured by pupillometry and subjective ratings of liking. Whether this optimum is mediated by individual differences would also be investigated.

5.6.3 Multiscale synchrony

The synchrony-bonding effect usually refers to synchrony of action at the timescale of seconds. Most human movement occurs within this timescale, e.g., walking at 0.83-1.95 Hz (Nilsson & Thorstensson, 1987). For the perceiver, this means processing roughly one stimulus per second, and making predictions about future stimuli that would occur approximately every second. However, similar principles could be applied at different timescales.

In music perception, rhythm at the timescale of seconds becomes pitch at the timescale of milliseconds; at slower speeds, auditory events are perceived by the brain but, at faster speeds, they are processed in the cochlea as pitch.[22] Therefore, harmony could be considered analogous to synchrony but at a faster timescale. There has been comparatively little research in the social bonding effects of harmony, although there is reason to believe that the evolution of the human singing voice may have been shaped by a need to harmonise (Bannan et al., 2022). Some have suggested that, in the domains of both pitch and rhythm, simple ratios may indicate that a stimulus has a biological origin (Bowling & Purves, 2015), so there could have

[22] See the following for a demonstration of this effect: https://dantepfer.com/blog/?p=277

been a selection pressure for being able to identify simple ratio stimuli at either timescale.[23] If low-level perceptual processes are involved in the synchrony-bonding effect, it's possible that the same or similar processes may be involved in a theoretical harmony-bonding effect, however this remains as purely speculation to be tested.

While rhythm at very fast timescales becomes pitch, rhythm at slower timescales becomes a calendar. It has already been noted that synchronised behaviour is an important feature of many rituals (Gelfand et al., 2020). However, rituals themselves may also represent a form of synchrony if they occur at regular time intervals. The doctrinal mode rituals described by Atkinson and Whitehouse (2011) are defined by their temporal regularity, as opposed to imagistic rituals which are more emotionally intense but less frequent. The regularity of the doctrinal ritual mode could be treated as a form of synchrony from a predictive processing perspective. An agent may predict that their social group will meet for an event every Friday evening, and every Friday evening they would have their prediction fulfilled. Three scales of synchrony could be engaged at once in this way, if the weekly ritual also involved coordinated pitch and rhythm. Future research may be able to explore whether ease of processing through the fulfillment of temporal predictions may operate across multiple timescales. These predictive processes may engage different regions of the brain, but still may fundamentally represent the same inferential process.

5.6.3.1 Alignment across domains

A further expansion of the role of synchrony (across timescales) could be considered in the multiple forms of interpersonal alignment suggested by Shamay-Tsoory (2019): behavioural, cognitive and emotional. It has long been noted that behavioural alignment (i.e., synchrony and imitation) may indicate similar internal states (Scheflen, 1964). As discussed,

[23] This could be particularly relevant for humans, given the human bipedal gait has a 2:1 rhythm ratio (Larsson et al., 2019). The existing research on the role of gait in rhythm perception remains highly speculative, however.

the hidden states of other agents can only be inferred from observing behaviour (Friston & Frith, 2015a). Behavioural states are observable, and can be aligned through synchronised behaviour, but cognitive and emotional states remain hidden.

Engaging in synchrony may be a means by which two agents can either access each other's hidden states or cause their hidden states to converge (Friston & Frith, 2015a). Given that coordination in general is easier between agents that share common ground (Vasil et al., 2020a), it may be possible to consider an agent's cognitive or emotional state as a predictor for successful synchrony. Aligning behaviour and cognition at multiple timescales (e.g., dancing at a ritual setting, with shared knowledge of the ritual) may lead to greater success in achieving synchrony. Conversely, further studies could investigate whether degree of synchrony predicts how accurately a participant can rate another's current cognitive or emotional state.

Forcing a misalignment in cognitive[24] states may make it harder to coordinate. For example, circadian rhythm influences cognitive abilities throughout the cycle of a day (Curtis et al., 2014). Two people at different stages in their circadian rhythm (through time-zone differences or shift work) may find their efforts impaired if they try to coordinate, as their cognitive states would be less mutually predictable. Such an issue would have been unlikely to occur before the invention of electric lights (Casiraghi et al., 2020), but it could now be commonplace. Research is starting to explore the increased cognitive load that accompanies videoconferencing due to the limited non-verbal cues available (Bailenson, 2021). However, this research does not fully consider the inherent temporal delay in videocalls that would

[24] Shamay-Tsoory and others (2019) do not include bodily alignment in their three factors of alignment but it could also be relevant here. The example of circadian rhythm is as much physiological as it is cognitive. Another example could be in preferred walking speed or leg length (Nilsson & Thorstensson, 1987). If two people have different natural walking speeds it would probably be more difficult for them to spontaneously synchronise their steps. Social factors are very important in running performance (Davis et al., 2021), yet such considerations may subconsciously shape social preferences when choosing a running partner.

disrupt the rhythm of conversation, nor the potential effects of a mismatched cognitive rhythm when interacting with someone in a different time-zone.

This conceptualisation of temporal coordination over multiple timescales between two (or more) mutually inferring agents that may be aligned across the domains of behaviour, emotion and cognition provides a rich framework for further experimentation. Regardless of the specific conditions, the fluency through synchrony theory would suggest that any agent would prefer to interact with another who is more predictable to them due to alignment across timescales and domains.

5.6.4 Multilevel selection

The synchrony-bonding effect may have provided selection pressure for the evolution of rhythmic abilities along with the evolution of music as a form of communication (Savage et al., 2021). Criticism of this stance cites its reliance on group selection (Mehr et al., 2021). However, the account presented here need not rely upon groups. The fluency through synchrony theory makes no particular assertions as to the selection pressures for rhythmic abilities. Crucially, processing fluency is a general principle that operates across perceptual domains, based upon the brain's optimisation principle to reduce energy costs (Reber et al., 2004) and did not evolve specifically for synchrony.

Processing fluency may not even be restricted to humans, although there is little research on non-human animals to suggest this. It remains to be fully explained exactly how humans moved from bonding through social grooming to synchrony (Launay et al., 2016). Based upon the fluency through synchrony theory, one possible account would be that early humans leveraged pre-existing perceptual processes to generate social bonding effects. If humans generally prefer simple, predictable stimuli, then there may have been pressure for other humans to make themselves simple and predictable to gain social favour.

It is perhaps pertinent that grooming has its own inherent rhythm (della Longa et al., 2020). Noting that rhythms originating from biological sources tend to have a simple structure (Bowling & Purves, 2015), this may mark the tactile sensation of being groomed as originating from a conspecific. Some research has suggested that the regular pulse of a musical beat stimulates the C-tactile system at a similar rate to grooming, through induced movement in the listener (Dunbar et al., 2021).[25] Synchronised movement does appear to stimulate the same endorphin reward systems as social grooming (Tarr et al., 2014), and the brain's reward systems are also active in response to processing fluency (Winkielman et al., 2003). There is still much to be understood about the precise neurohormonal pathways involved, however, and future research would have to expand upon the biological basis for the fluency through synchrony theory.

It is beyond the scope of this thesis to outline the entire evolutionary history of synchrony and human rhythmic abilities, although speculation can be made in light of the findings here. If processing fluency was leveraged by early humans, to supplant grooming as the primary bonding mechanism, this could have occurred at either the group or individual level according to a multi-level selection model (Okasha, 2006). Groups that were able to temporally coordinate behaviours may have been able to reduce processing load for each other, thus making the whole group more cognitively efficient during the task. However, an individual with rhythmic abilities superior to their conspecifics may find themselves at an advantage, as they would be better able to reduce processing load for others, thus creating a positive impression of themselves. Given the basic nature of processing fluency, the ability to perceive synchrony probably preceded the ability for rhythmic entrainment. Any individual in

[25] It is also worth noting that, in addition to being activated while nodding to music, the vestibular system aids in constructing concepts of the self (Lenggenhager & Lopez, 2015).

possession of the latter would thus find themselves with the ability to manipulate another's perception. Being able to reduce processing load for others may have given an individual social advantages within a group, as much as it would help the group as a whole.

5.6.5 The future of synchrony

Finally, a comment should be made on the role of synchrony within an increasingly desynchronised world. A process of de-ritualisation in most Western cultures has resulted in fewer opportunities for synchronised action (Garfinkel, 2010; Nietzsche, 1908; Veissiere, 2021). This was well underway since the invention of technologies that enable on-demand media such as books, recorded music, video, and the world wide web. This contrasts with the social consumption of media that would have defined human social interactions for most of our evolutionary history: storytelling, group music making, and dance. Before the invention of recorded music, listening to music required live performance and a temporally aligned experience with other people.

The social distancing required by the COVID-19 pandemic brought new challenges for synchronised behaviour. Many in the performing arts were forced to cancel in-person performances, and both musicians and their audiences turned to internet videos and livestreamed concerts (Greenberg et al., 2021). Online musical activity poses particular challenges, due to the inherent delay in video calling, making live performance between multiple people in different locations almost impossible. Nevertheless, many musicians found ways of producing content during the pandemic (Hansen et al., 2021).

This also provided a rich new study area to test the social effects of online events. Multiple studies have now found that livestream concerts can facilitate feelings of social connectedness, when compared with watching pre-recorded concerts (Onderdijk et al., 2021; Swarbrick et al., 2021). Even without the tight, temporal synchrony of in-person interactions,

just knowing that others are watching the same video at roughly the same time seems to produce some effect. Furthermore, many people found themselves listening to more music during lockdowns (Carlson et al., 2021), possibly because listening to music may be perceived as a social-surrogate (Schäfer et al., 2020), a kind of social interaction with the performer, even if that performer is recorded.

There is much more research to be done in this regard, but it highlights the resilience of people's desire for some level of synchrony. This is important to recognise, as even before the COVID-19 pandemic, many countries such as the UK were already dealing with a rapid increase in loneliness (UK Department for Digital, Culture, 2018). It has also been noted that the general pro-sociality following synchronised behaviour may lead to people making more sustainable and socially conscious choices (Soundy et al., 2022). Given the challenges facing society in 2022, it has never been more crucial to understand the ways in which humans create and maintain social bonds.

5.7 Final conclusions

In this thesis I have outlined a novel mechanism to explain the synchrony-bonding effect. Although the synchrony-bonding effect has been observed in numerous studies, there is disagreement as to the precise mechanisms underpinning the effect. Most of the proposed mechanisms draw upon higher-order reasoning about the interaction, but very few have considered the perceptual experience of engaging with synchrony. Based upon the general principles of processing fluency, I suggest that people generally like synchrony, and by extension those with which they are synchronised, because synchrony is easier to perceive than non-synchrony.

Chapter 1 outlined the existing literature on processing fluency and how it relates to synchrony. Chapter 2 provided empirical evidence that synchrony is easier to perceive than

non-synchrony (Study 1). Chapter 3 built upon this with a rhythmic tapping task, finding that synchrony is easier to participate in than non-synchrony (Study 2). Chapter 4 used a variation of the task from Study 2 to build a mediation model in which the social bonding effects of synchrony are explained by the mediating factor of perceived task difficulty (Study 3). This final chapter has discussed the evidence from these empirical studies in light of the broader theoretical context.

I have attempted to integrate the proposed fluency through synchrony theory with some of the existing theories explaining the synchrony-bonding effect. Many of these theories may be fundamentally compatible but operating at different levels of explanation. Crucially, the processing fluency effects observed here are likely subconscious and may be open to different conscious interpretations depending upon context. I suggest that processing fluency may play a role in directing social attention during, and on the motivation to perform, synchronised action. Furthermore, the experience of self-other overlap often reported in synchrony-bonding studies could reflect either a perceptual or predictive processing experience in which one may incorporate another into their sense of self through perceptual streaming or self-prediction, both of which fundamentally reflect a reduction in processing load. Either the other is easier to process because their actions are perceptually indistinguishable from one's own, or because they are predictable from one's own.

The fluency through synchrony theory, as outlined in this thesis, opens multiple avenues of further enquiry. This theory considers behavioural synchrony as a system with two mutually inferring agents that are attempting to reduce each other's processing load. With this as a starting point, future research could investigate the optimum amount of complexity in any coordinated behaviour to produce social bonding outcomes, as well as some broader contextual and interpersonal factors. It allows one to consider synchrony across multiple timescales. It

also may contribute to ongoing debates about the evolutionary function of musical behaviour, and the selection pressures for the abilities that support it.

Ultimately, this thesis puts forward a case that synchrony is easier to process than non-synchrony and that when a dyad acts in synchrony the partners make themselves easier for each other to perceive. This reduced processing load is inherently rewarding, leading both partners in the dyad to feel positively about their interaction partner. It frames the act of entrainment as a process of making oneself predictable to reduce another's processing load. As the dancer, Shane McCarthy states 'I dance with someone to make them feel good about themselves' (pers. com., 2021). We may not be able to reduce all the unpredictability in our increasingly complex world through synchrony, but it cannot hurt to try.

References

Abernethy, B. (1988). Dual-task methodology and motor skills research: Some applications and methodological constraints. *Journal of Human Movement Studies*, *14*(3), 101–132. https://espace.library.uq.edu.au/view/UQ:276056

Agrawal, Y., Carlson, E., Toiviainen, P., & Alluri, V. (2022). Decoding Individual differences and musical preference via music-induced movement. *Scientific Reports*, *12*(1), 1–14. https://doi.org/10.1038/s41598-022-06466-3

Alain, C., & Izenberg, A. (2003). Effects of attentional load on auditory scene analysis. *Journal of Cognitive Neuroscience*, *15*(7), 1063–1073.

Alnæs, D., Sneve, M. H., Espeseth, T., Endestad, T., van de Pavert, S. H. P., & Laeng, B. (2014). Pupil size signals mental effort deployed during multiple object tracking and predicts brain activity in the dorsal attention network and the locus coeruleus. *Journal of Vision*, *14*(4). https://doi.org/10.1167/14.4.1

Amano, K., Edwards, M., Badcock, D. R., & Nishida, S. (2009). Adaptive pooling of visual motion signals by the human visual system revealed with a novel multi-element stimulus. *Journal of Vision*, *9*(3), 4–4. https://doi.org/10.1167/9.3.4

Apps, M. A. J., & Tsakiris, M. (2014). The free-energy self: A predictive coding account of self-recognition. *Neuroscience and Biobehavioral Reviews*, *41*, 85–97. https://doi.org/10.1016/j.neubiorev.2013.01.029

Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of Other in the Self Scale and the Structure of Interpersonal Closeness. *Journal of Personality and Social Psychology*, *63*(4), 596–612. https://doi.org/10.1037/0022-3514.63.4.596

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. https://doi.org/10.1016/j.evolhumbehav.2010.09.002

Attard-Johnson, J., Ó Ciardha, C., & Bindemann, M. (2019). Comparing methods for the analysis of pupillary response. *Behavior Research Methods*, *51*(1), 83–95. https://doi.org/10.3758/s13428-018-1108-6

Atwood, S., Schachner, A., & Mehr, S. A. (2020). Expectancy effects threaten the inferential validity of synchrony-prosociality research. *PsyArXiv*. 10.31234/osf.io/zjy8u

Avitabile, D., Słowiński, P., Bardy, B., & Tsaneva-Atanasova, K. (2016). Beyond in-phase and anti-phase coordination in a model of joint action. *Biological Cybernetics*, *110*(2–3), 201–216. https://doi.org/10.1007/s00422-016-0691-9

Bailenson, J. N. (2021). Nonverbal overload: A theoretical argument for the causes of Zoom fatigue. *Technology, Mind, and Behavior*, *2*(1). https://doi.org/10.1037/tmb0000030

Bamford, J. S. (2017). *Social bonding happens in time: Interpersonal synchronisation in the silent disco* [Masters, University of Jyväskylä]. https://doi.org/10.13140/RG.2.2.13434.54723

Bamford, J. S., Burger, B., & Toiviainen, P. (2016). Are we dancing to the same beat? empathy and interpersonal synchronisation in the silent disco. In B. Burger, J. S. Bamford, & E. Carlson (Eds.), *The 9th International Conference of Students of Systematic Musicology (SysMus16): Programme, Abstracts & Proceedings* (Issue June). University of Jyväskylä.

Bamford, J. S., & Davidson, J. W. (2019). Trait Empathy associated with Agreeableness and rhythmic entrainment in a spontaneous movement to music task: Preliminary exploratory investigations. *Musicae Scientiae*, *23*(1), 5–24. https://doi.org/10.1177/1029864917701536

Bannan, N., Bamford, J. S., & Dunbar, R. I. M. (2022). *The evolution of gender dimorphism in the human voice: the role of octave equivalence*. Current Anthropology. https://ora.ox.ac.uk/objects/uuid:12819bce-6066-4ee0-841e-c6ac831afdac

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin*, *91*(2), 276–292. https://doi.org/10.1037/0033-2909.91.2.276

Belke, B., Leder, H., Strobach, T., & Carbon, C. C. (2010). Cognitive Fluency: High-Level Processing Dynamics in Art Appreciation. *Psychology of Aesthetics, Creativity, and the Arts*, *4*(4), 214–222. https://doi.org/10.1037/a0019648

Bendixen, A. (2014). Predictability effects in auditory scene analysis: A review. *Frontiers in Neuroscience*, *8*(8 MAR), 1–16. https://doi.org/10.3389/fnins.2014.00060

Bernieri, F. J., Reznick, J. S., & Rosenthal, R. (1988). Synchrony, pseudosynchrony, and dissynchrony: measuring the entrainment process in mother-infant interactions. *Journal of Personality and Social Psychology*, *54*(2), 243–253.

Bernieri, F. J., & Rosenthal, R. (1991). Interpersonal coordination: Behavior matching and interactional synchrony. In R. S. Feldman & B. Rimé (Eds.), *Fundamentals of nonverbal behavior* (Issue February, pp. 401–432). Cambridge University Press. https://psycnet.apa.org/record/1991-98576-011

Berridge, K. C., & Kringelbach, M. L. (2008). Affective neuroscience of pleasure: Reward in humans and animals. *Psychopharmacology*, *199*(3), 457–480. https://doi.org/10.1007/s00213-008-1099-6

Berridge, K. C., & Kringelbach, M. L. (2015). Pleasure Systems in the Brain. *Neuron*, *86*(3), 646–664. https://doi.org/10.1016/j.neuron.2015.02.018

Bianco, R., Ptasczynski, L. E., Omigie, D., Roberta, B., Esther, P. L., & Diana, O. (2019). Pupil responses to pitch deviants reflect predictability of melodic sequences. *BioRxiv*, 693382. https://doi.org/10.1101/693382

Bidelman, G. M., & Krishnan, A. (2009). *Behavioral/Systems/Cognitive Neural Correlates of Consonance, Dissonance, and the Hierarchy of Musical Pitch in the Human Brainstem*. https://doi.org/10.1523/JNEUROSCI.3900-09.2009

Bieńkiewicz, M. M. N., Smykovskyi, A., Olugbade, T., Janaqi, S., Camurri, A., Bianchi-Berthouze, N., Björkman, M., & Bardy, B. G. (2021). Bridging the gap between emotion and joint action. *Neuroscience & Biobehavioral Reviews*. https://doi.org/10.1016/j.neubiorev.2021.08.014

Bishop, L., Jensenius, A. R., & Laeng, B. (2021). Musical and Bodily Predictors of Mental Effort in String Quartet Music: An Ecological Pupillometry Study of Performers and Listeners. *Frontiers in Psychology*, *12*(June). https://doi.org/10.3389/fpsyg.2021.653021

Bispham, J. C. (2018). *Music's Evolutionary 'Design Features'* [Doctoral]. University of Cambridge.

Bornstein, R. F., & D'Agostino, P. R. (1992). Stimulus Recognition and the Mere Exposure Effect. *Journal of Personality and Social Psychology*, *63*(4), 545–552. https://doi.org/10.1037/0022-3514.63.4.545

Botvinick, M., & Cohen, J. (1998). Rubber hands 'feel' touch that eyes see. *Nature*, *391*(6669), 756–756. https://doi.org/10.1038/35784

Bouwer, F. L., Nityananda, V., Rouse, A. A., & ten Cate, C. (2021). Rhythmic abilities in humans and non-human animals: a review and recommendations from a methodological perspective. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *376*(1835), 20200335. https://doi.org/10.1098/rstb.2020.0335

Bowling, D. L., Graf Ancochea, P., Hove, M. J., & Fitch, W. T. (2019). Pupillometry of Groove: Evidence for Noradrenergic Arousal in the Link Between Music and Movement. *Frontiers in Neuroscience*, *12*(January), 1039. https://doi.org/10.3389/fnins.2018.01039

Bowling, D. L., & Purves, D. (2015). A biological rationale for musical consonance. In *Proceedings of the National Academy of Sciences of the United States of America* (Vol. 112, Issue 36, pp. 11155–11160). National Academy of Sciences. https://doi.org/10.1073/pnas.1505768112

Bregman, A. S. (1990). *Auditory Scene Analysis*. MIT Press.

Bregman, A. S., & Campbell, J. (1971). Primary auditory stream segregation and perception of order in rapid sequences of tones. *Journal of Experimental Psychology*, *89*(2), 244–249. https://doi.org/10.1037/h0031163

Brown, H., Adams, R. A., Parees, I., Edwards, M., & Friston, K. J. (2013). *Active inference, sensory attenuation and illusions*. https://doi.org/10.1007/s10339-013-0571-3

Burger, B., Thompson, M. R., Luck, G., Saarikallio, S. H., & Toiviainen, P. (2014). Hunting for the beat in the body: On period and phase locking in music-induced movement. *Frontiers in Human Neuroscience*, *8*(November), 1–16. https://doi.org/10.3389/fnhum.2014.00903

Burger, B., & Toiviainen, P. (2013). MoCap Toolbox - a Matlab toolbox for computational analysis of movement data. In R. Bresin (Ed.), *Proceedings of the Sound and Music Computing Conference* (pp. 172–178). Logos Verlag Berlin.

Butler, L. T., & Berry, D. C. (2004). Understanding the relationship between repetition priming and mere exposure. *British Journal of Psychology*, *95*(4), 467–487. https://doi.org/10.1348/0007126042369776

Carlson, E., Burger, B., London, J., Thompson, M. R., & Toiviainen, P. (2016). Conscientiousness and Extraversion relate to responsiveness to tempo in dance. *Human Movement Science*, *49*, 315–325. https://doi.org/10.1016/j.humov.2016.08.006

Carlson, E., Saari, P., Burger, B., & Toiviainen, P. (2020). Dance to your own drum: Identification of musical genre and individual dancer from motion capture using machine learning. *Journal of New Music Research*, *0*(0), 1–16. https://doi.org/10.1080/09298215.2020.1711778

Carlson, E., Wilson, J., Baltazar, M., Duman, D., Peltola, H. R., Toiviainen, P., & Saarikallio, S. (2021). The Role of Music in Everyday Life During the First Wave of the Coronavirus Pandemic: A Mixed-Methods Exploratory Study. *Frontiers in Psychology*, *12*(May), 1–15. https://doi.org/10.3389/fpsyg.2021.647756

Casiraghi, L. P., Plano, S. A., Fernández-Duque, E., Valeggia, C., Golombek, D. A., & de la Iglesia, H. O. (2020). Access to electric light is associated with delays of the dim-light melatonin onset in a traditionally hunter-gatherer Toba/Qom community. *Journal of Pineal Research*, *69*(4). https://doi.org/10.1111/jpi.12689

Chapman, S. (1968). Catching a Baseball. *American Journal of Physics*, *36*(10), 868–870. https://doi.org/10.1119/1.1974297

Chen, A., Tang, D., & Chen, X. (2013). Training Reveals the Sources of Stroop and Flanker Interference Effects. *PLoS ONE*, *8*(10). https://doi.org/10.1371/journal.pone.0076580

Chirico, A., Serino, S., Cipresso, P., Gaggioli, A., & Riva, G. (2015). When music "flows". State and trait in musical performance, composition and listening: a systematic review. *Frontiers in Psychology*, *6*, 906. https://doi.org/10.3389/fpsyg.2015.00906

Chmiel, A., & Schubert, E. (2017). Back to the inverted-U for music preference: A review of the literature. *Psychology of Music*, *45*(6), 886–909. https://doi.org/10.1177/0305735617697507

Cirelli, L. K. (2018). How interpersonal synchrony facilitates early prosocial behavior. *Current Opinion in Psychology*, *20*, 35–39. https://doi.org/10.1016/j.copsyc.2017.08.009

Cirelli, L. K., Einarson, K. M., & Trainor, L. J. (2014). Interpersonal synchrony increases prosocial behavior in infants. *Developmental Science*, *17*(6), 1003–1011. https://doi.org/10.1111/desc.12193

Clark, A. (2015). Embodied Prediction. *Open MIND*, *7*, 1–21. https://doi.org/10.15502/9783958570115

Clark, A. (2018). A nice surprise? Predictive processing and the active pursuit of novelty. *Phenomenology and the Cognitive Sciences*, *17*(3), 521–534. https://doi.org/10.1007/s11097-017-9525-z

Clayton, M. (2016). The social and personal functions of music in cross-cultural perspective. In S. Hallam, I. Cross, & M. Thaut (Eds.), *The Oxford Handbook of Music Psychology* (2nd ed., p. 950). 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), 106–108. https://doi.org/10.1098/rsbl.2009.0670

Cohen, E., Mundry, R., & Kirschner, S. (2014). Religion, synchrony, and cooperation. *Religion, Brain and Behavior*, *4*(1), 20–30. https://doi.org/10.1080/2153599x.2012.741075

Costa, V. D., & Rudebeck, P. H. (2016). More than meets the eye: the relationship between pupil size and Locus Coerulus activity. *Neuron*, *89*(1), 8–10. https://doi.org/10.1016/j.physbeh.2017.03.040

Cross, L., Atherton, G., & Sebanz, N. (2021). Intentional synchronisation affects automatic imitation and source memory. *Scientific Reports*, *11*(573). https://doi.org/10.1038/s41598-020-79796-9

Crump, M. J. C., McDonnell, J. V., & Gureckis, T. M. (2013). Evaluating Amazon's Mechanical Turk as a Tool for Experimental Behavioral Research. *PLoS ONE*, *8*(3). https://doi.org/10.1371/journal.pone.0057410

Csikszentmihalyi, M., & Csikszentmihalyi, I. S. (1988). *Optimal experience: psychological studies of flow in consciousness*. Cambridge University Press.

Cupchik, G. C., Vartanian, O., Crawley, A., & Mikulis, D. J. (2009). Viewing artworks: Contributions of cognitive control and perceptual facilitation to aesthetic experience. *Brain and Cognition*, *70*(1), 84–91. https://doi.org/10.1016/j.bandc.2009.01.003

Curtis, J., Burkley, E., & Burkley, M. (2014). The Rhythm Is Gonna Get You: The Influence of Circadian Rhythm Synchrony on Self-Control Outcomes. *Social and Personality Psychology Compass*, *8*, 609–625. https://doi.org/10.1111/spc3.12136

Davis, A. J., MacCarron, P., & Cohen, E. (2021). Social reward and support effects on exercise experiences and performance: Evidence from parkrun. *PLoS ONE*, *16*(9 September). https://doi.org/10.1371/JOURNAL.PONE.0256546

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods*, *47*(1), 1–12. https://doi.org/10.3758/s13428-014-0458-y

De Valois, R. L., & De Valois, K. K. (1991). Vernier acuity with stationary moving Gabors. *Vision Research*, *31*(9), 1619–1626. https://doi.org/10.1016/0042-6989(91)90138-U

della Longa, L., Filippetti, M. L., Dragovic, D., & Farroni, T. (2020). Synchrony of Caresses: Does Affective Touch Help Infants to Detect Body-Related Visual–Tactile Synchrony? *Frontiers in Psychology*, *10*(January), 1–10. https://doi.org/10.3389/fpsyg.2019.02944

Demos, A. P., Chaffin, R., Begosh, K. T., Daniels, J. R., & Marsh, K. L. (2012). Rocking to the beat: Effects of music and partner's movements on spontaneous interpersonal coordination. *Journal of Experimental Psychology: General*, *141*(1), 49–53. https://doi.org/10.1037/a0023843

Diehl, M. (1990). The minimal group paradigm: Theoretical explanations and empirical findings. *European Review of Social Psychology*, *1*(1), 263–292. https://doi.org/10.1080/14792779108401864

Dolležal, L.-V., Brechmann, A., Deike, S., & Klump, G. M. (2014). Evaluating auditory stream segregation of SAM tone sequences by subjective and objective psychoacoustical tasks, and brain activity. *Frontiers in Neuroscience*, *8*, 119. https://doi.org/10.3389/fnins.2014.00119

Duchowski, A. T., Krejtz, K., Gehrer, N. A., Bafna, T., & Baekgaard, P. (2020). The Low / High Index of Pupillary Activity. *CHI*, 1–12.

Duchowski, A. T., Krejtz, K., Krejtz, I., Biele, C., Niedzielska, A., Kiefer, P., Raubal, M., & Giannopoulos, I. (2018). The Index of Pupillary Activity Measuring Cognitive Load vis-à-vis Task Difficulty with Pupil Oscillation. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3173574.3173856

Dunbar, R. I. M. (1991). Functional Significance of Social Grooming in Primates. *Folia Primatologica*, *57*(3), 121–131. https://doi.org/10.1159/000156574

Dunbar, R. I. M. (2009). The social brain hypothesis and its implications for social evolution. *Annals of Human Biology*, *36*(5), 562–572. https://doi.org/10.1080/03014460902960289

Dunbar, R. I. M. (2010). The social role of touch in humans and primates: Behavioural function and neurobiological mechanisms. *Neuroscience and Biobehavioral Reviews*, *34*(2), 260–268. https://doi.org/10.1016/j.neubiorev.2008.07.001

Dunbar, R. I. M. (2012a). Bridging the bonding gap: The transition from primates to humans. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1597), 1837–1846. https://doi.org/10.1098/rstb.2011.0217

Dunbar, R. I. M. (2012b). On the evolutionary function of song and dance. In N. Bannan (Ed.), *Music, Language & Human Evolution* (p. 201). Oxford University Press.

Dunbar, R. I. M., Pearce, E., Tarr, B., Makdani, A., Bamford, J. S., Smith, S., & Mcglone, F. (2021). Cochlear SGN neurons elevate pain thresholds in response to music. *Scientific Reports*, 1–8. https://doi.org/10.1038/s41598-021-93969-0

Durkheim, E. (1995). *The Elementary Forms of Religious Life*. The Free Press.

Dutilh, G., Wagenmakers, E.-J., Visser, I., & Van Der Maas, H. L. J. (2011). A Phase Transition Model for the Speed-Accuracy Trade-Off in Response Time Experiments. *Cognitive Science*, *35*(2), 211–250. https://doi.org/10.1111/j.1551-6709.2010.01147.x

Elhilali, M., Ma, L., Micheyl, C., Oxenham, A. J., & Shamma, S. A. (2009). Temporal Coherence in the Perceptual Organization and Cortical Representation of Auditory Scenes. *Neuron*, *61*(2), 317–329. https://doi.org/10.1016/j.neuron.2008.12.005

Ellis, D. P. W. (1999). Using knowledge to organize sound: The prediction-driven approach to computational auditory scene analysis and its application to speech/nonspeech mixtures. *Speech Communication*, *27*(3), 281–298. https://doi.org/10.1016/S0167-6393(98)00083-1

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*(1), 143–149.

Fairhurst, M. T., Janata, P., & Keller, P. E. (2013). Being and Feeling in Sync with an Adaptive Virtual Partner: Brain Mechanisms Underlying Dynamic Cooperativity. *Cerebral Cortex*, *23*(11), 2592–2600. https://doi.org/10.1093/cercor/bhs243

Fayn, K., MacCann, C., Tiliopoulos, N., & Silvia, P. J. (2015). Aesthetic Emotions and Aesthetic People: Openness Predicts Sensitivity to Novelty in the Experiences of Interest and Pleasure. *Frontiers in Psychology*, *6*, 1–11. https://doi.org/10.3389/fpsyg.2015.01877

Fitch, W. T. (2013). Rhythmic cognition in humans and animals: distinguishing meter and pulse perception. *Frontiers in Systems Neuroscience*, *7*. https://doi.org/10.3389/fnsys.2013.00068

Fitch, W. T. (2015). Four principles of bio-musicology. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1664). https://doi.org/10.1098/rstb.2014.0091

Fiveash, A., McArthur, G., & Thompson, W. F. (2018). Syntactic and non-syntactic sources of interference by music on language processing. *Scientific Reports*, *8*(1), 1–15. https://doi.org/10.1038/s41598-018-36076-x

Forster, M., Gerger, G., & Leder, H. (2015). Everything's relative? Relative differences in processing fluency and the effects on liking. *PLoS ONE*, *10*(8). https://doi.org/10.1371/journal.pone.0135944

Forsythe, A., Nadal, M., Sheehy, N., Cela-Conde, C. J., & Sawey, M. (2011). Predicting beauty: Fractal dimension and visual complexity in art. *British Journal of Psychology*, *102*(1), 49–70. https://doi.org/10.1348/000712610X498958

Fraser, T., Hew, A., Crooke, D., & Davidson, J. W. (2021). *"Music Has No Borders": An Exploratory Study of Audience Engagement With YouTube Music Broadcasts During COVID-19*. *12*(July), 1–17. https://doi.org/10.3389/fpsyg.2021.643893

Friston, K. J., & Frith, C. (2015a). A Duet for one. *Consciousness and Cognition*, *36*, 390–405. https://doi.org/10.1016/j.concog.2014.12.003

Friston, K. J., & Frith, C. D. (2015b). Active inference, Communication and hermeneutics. *Cortex*, *68*(Kelso 2012), 129–143. https://doi.org/10.1016/j.cortex.2015.03.025

Friston, K. J., & Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521), 1211–1221. https://doi.org/10.1098/rstb.2008.0300

Friston, K. J., Thornton, C., & Clark, A. (2012). Free-Energy Minimization and the Dark-Room Problem. *Frontiers in Psychology*, *3*(MAY), 1–7. https://doi.org/10.3389/fpsyg.2012.00130

Gao, C. J., & Gozli, D. G. (2021). Are self-caused distractors easier to ignore? Experiments with the flanker task. *Attention, Perception, and Psychophysics*, *83*(2), 853–865. https://doi.org/10.3758/s13414-020-02170-4

Garfinkel, Y. (2010). Dance in prehistoric Europe. *Documenta Praehistorica*, *37*(1), 205–214. https://doi.org/10.4312/dp.37.18

Gelfand, M. J., Caluori, N., Jackson, J. C., & Taylor, M. K. (2020). *The cultural evolutionary trade-off of ritualistic synchrony*. https://doi.org/10.1098/rstb.2019.0432

Goebl, W., & Guggenberger, D. (2015). Tappingfriend - an interactive science exhibit for experiencing synchronicity with real and artificial partners. *Proceedings of the Third Vienna Talk on Music Acoustics*, 227–230.

Greenberg, D. M., Decety, J., & Gordon, I. (2021). The Social Neuroscience of Music: Understanding the Social Brain Through Human Song. *American Psychologist*, *76*(7), 1172–1185. https://doi.org/10.1037/amp0000819

Greifeneder, R., Alt, A., Bottenberg, K., Seele, T., Zelt, S., & Wagener, D. (2010). On Writing Legibly. *Social Psychological and Personality Science*, *1*(3), 230–237. https://doi.org/10.1177/1948550610368434

Hampson, R. E., Opris, I., & Deadwyler, S. A. (2010). Neural correlates of fast pupil dilation in nonhuman primates: Relation to behavioral performance and cognitive workload. *Behavioural Brain Research*, *212*(1), 1–11. https://doi.org/10.1016/j.bbr.2010.03.011

Hansen, N. Chr., Treider, J. M. G., Swarbrick, D., Bamford, J. S., Wilson, J., & Vuoskoski, J. K. (2021). A Crowd-Sourced Database of Coronamusic: Documenting Online Making and Sharing of Music During the COVID-19 Pandemic. *Frontiers in Psychology*, *12*(June), 1–9. https://doi.org/10.3389/fpsyg.2021.684083

Harvey, A. R. (2020). Links Between the Neurobiology of Oxytocin and Human Musicality. *Frontiers in Human Neuroscience*, *14*(August), 1–19. https://doi.org/10.3389/fnhum.2020.00350

Hasegawa, R. (2021). Timbre as Harmony—Harmony as Timbre. *The Oxford Handbook of Timbre*, 524–551. https://doi.org/10.1093/OXFORDHB/9780190637224.013.11

Hesp, C., Smith, R., Parr, T., Allen, M., Friston, K. J., & Ramstead, M. (2019). *Deeply Felt Affect: The Emergence of Valence in Deep Active Inference*. *August 2020*. https://doi.org/10.31234/osf.io/62pfd

Houtsma, A. J. M. (1997). Pitch and timbre: Definition, meaning and use. *Journal of New Music Research*, *26*(2), 104–115. https://doi.org/10.1080/09298219708570720

Hove, M. J., & Risen, J. L. (2009). It's all in the timing: Interpersonal synchrony increases affiliation. *Social Cognition*, *27*(6), 949–960. https://doi.org/10.1521/soco.2009.27.6.949

Huang, H.-J., & Mercer, V. S. (2001). Dual-task methodology: applications in studies of cognitive and motor performance in adults and children. *Pediatric Physical Therapy*, *13*(3), 133–140.

Hughes, R. W., Hurlstone, M. J., Marsh, J. E., Jones, D. M., & Vachon, F. (2013). Cognitive control of auditory distraction: Impact of task difficulty, foreknowledge, and working memory capacity supports duplex-mechanism account. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(2), 539–553. https://doi.org/10.1037/a0029064

Hukin, R. W., & Darwin, C. J. (1995). Comparison of the effect of onset asynchrony on auditory grouping in pitch matching and vowel identification. *Perception & Psychophysics*, *57*(2), 191–196. https://doi.org/10.3758/BF03206505

Hunter, P. G., & Schellenberg, E. G. (2011). Interactive effects of personality and frequency of exposure on liking for music. *Personality and Individual Differences*, *50*(2), 175–179. https://doi.org/10.1016/j.paid.2010.09.021

Jackson, S. A., & Marsh, H. W. (1996). Development and validation of a scale to measure optimal experience: The flow state scale. *Journal of Sport and Exercise Psychology*, *18*(1), 17–35. https://doi.org/10.1123/jsep.18.1.17

Jacob, R. J. K., & Karn, K. S. (2003). Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. In Hyona, Radach, & Deubel (Eds.), *The Mind's Eye* (pp. 573–605). Elsevier Science. https://www.sciencedirect.com/science/article/pii/B9780444510204500311

Jacoby, N., & McDermott, J. H. (2017). Integer Ratio Priors on Musical Rhythm Revealed Cross-culturally by Iterated Reproduction. *Current Biology*, *27*(3), 359–370. https://doi.org/10.1016/j.cub.2016.12.031

Jakesch, M., Leder, H., & Forster, M. (2013). Image Ambiguity and Fluency. *PLoS ONE*, *8*(9), e74084. https://doi.org/10.1371/journal.pone.0074084

JASP Team. (2020). *JASP* (0.14.1). https://jasp-stats.org/

Jiménez, J., Gómez, Á., Buhrmester, M. D., Vázquez, A., Whitehouse, H., & Swann, W. B. (2016). The Dynamic Identity Fusion Index: A New Continuous Measure of Identity Fusion for Web-Based Questionnaires. *Social Science Computer Review*, *34*(2), 215–228. https://doi.org/10.1177/0894439314566178

Johnston, W. A., Hawley, K. J., Plewe, S. H., Elliott, J. M. G., & DeWitt, M. J. (1990). Attention Capture by Novel Stimuli. *Journal of Experimental Psychology: General*, *119*(4), 397–411.

Joshi, S., Li, Y., Kalwani, R. M., & Gold, J. I. (2016). Relationships between Pupil Diameter and Neuronal Activity in the Locus Coeruleus, Colliculi, and Cingulate Cortex. *Neuron*, *89*(1), 1–14. https://doi.org/10.1016/j.neuron.2015.11.028

Kahl, S., & Kopp, S. (2018). A predictive processing model of perception and action for self-other distinction. *Frontiers in Psychology*, *9*(DEC), 2421. https://doi.org/10.3389/FPSYG.2018.02421/BIBTEX

Kassner, M., Patera, W., & Bulling, A. (2014). Pupil: an open source platform for pervasive eye tracking and mobile gaze-based interaction. *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*, 1151–1160. https://doi.org/10.1145/2638728.2641695

Kawakami, K., Friesen, J. P., Williams, A., Vingilis-Jaremko, L., Sidhu, D. M., Rodriguez-Bailón, R., Cañadas, E., & Hugenberg, K. (2021). Impact of perceived interpersonal similarity on attention to the eyes of same-race and other-race faces. *Cognitive Research: Principles and Implications*, *6*(1). https://doi.org/10.1186/s41235-021-00336-8

Keller, P. E. (2001). Attentional resource allocation in musical ensemble performance. *Psychology of Music*, *29*(1), 20–38. https://doi.org/10.1177/0305735601291003

Keller, P. E. (2013). Musical ensemble performance: A theoretical framework and empirical findings on interpersonal coordination. *Proceedings of the International Symposium on Performance Science 2013*, 271–285.

Keller, P. E., Novembre, G., & Hove, M. J. (2014). Rhythm in joint action: psychological and neurophysiological mechanisms for real-time interpersonal coordination. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1658), 20130394. https://doi.org/10.1098/rstb.2013.0394

Kelley, C. M., & Lindsay, D. S. (1993). Remembering mistaken for knowing: ease of retrieval as a basis for confidence in answers to general knowledge questions. *Journal of Memory and Language*, *32*, 1–24.

Kiefer, P., Giannopoulos, I., Duchowski, A., & Raubal, M. (2016). Measuring cognitive load for map tasks through pupil diameter. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *9927 LNCS*, 323–337. https://doi.org/10.1007/978-3-319-45738-3_21

Killingly, C., Lacherez, P., & Meuter, R. (2021). Singing in the brain: investigating the cognitive basis of earworms. *Music Perception*, *38*(5), 456–472. https://doi.org/10.1525/mp.2021.38.5.456

Kirschner, S., & Tomasello, M. (2009). Joint drumming: Social context facilitates synchronization in preschool children. *Journal of Experimental Child Psychology*, *102*(3), 299–314. https://doi.org/10.1016/j.jecp.2008.07.005

Knoblich, G., & Sebanz, N. (2006). The social nature of perception and action. *Current Directions in Psychological Science*, *15*(3), 99–104.

Koban, L., Ramamoorthy, A., & Konvalinka, I. (2019). Why do we fall into sync with others? Interpersonal synchronization and the brain's optimization principle. In *Social Neuroscience* (Vol. 14, Issue 1, pp. 1–9). Taylor and Francis Ltd. https://doi.org/10.1080/17470919.2017.1400463

Koelsch, S., Vuust, P., & Friston, K. J. (2019). Predictive Processes and the Peculiar Case of Music. *Trends in Cognitive Sciences*, *23*(1), 63–77. https://doi.org/10.1016/j.tics.2018.10.006

Konvalinka, I., Vuust, P., Roepstorff, A., & Frith, C. D. (2010). Follow you, follow me: Continuous mutual prediction and adaptation in joint tapping. *Quarterly Journal of Experimental Psychology*, *63*(11), 2220–2230. https://doi.org/10.1080/17470218.2010.497843

Koole, S. L., Tschacher, W., Butler, E., Dikker, S., & Wilderjans, T. (2020). In sync with your shrink: grounding psychotherapy in interpersonal synchrony. In J. P. Forgas, W. D. Crano, & K. Fiedler (Eds.), *Applications of Social Psychology* (1st ed.). Routledge.

Koriat, A., & Levy-Sadot, R. (2001). The combined contributions of the cue-familiarity and accessibility heuristics to feelings of knowing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(1), 34–53. https://doi.org/10.1037/0278-7393.27.1.34

Kornmeier, J., & Bach, M. (2012). Ambiguous figures - what happens in the brain when perception changes but not the stimulus. *Frontiers in Human Neuroscience*, *6*(MARCH 2012), 1–23. https://doi.org/10.3389/fnhum.2012.00051

Kotz, S. A., Ravignani, A., & Fitch, W. T. (2018). The Evolution of Rhythm Processing. In *Trends in Cognitive Sciences* (Vol. 22, Issue 10, pp. 896–910). Elsevier Ltd. https://doi.org/10.1016/j.tics.2018.08.002

Krause, A. E., Dimmock, J., Rebar, A. L., & Jackson, B. (2021). Music Listening Predicted Improved Life Satisfaction in University Students During Early Stages of the COVID-19 Pandemic. *Frontiers in Psychology*, *11*(January), 1–9. https://doi.org/10.3389/fpsyg.2020.631033

Krejtz, K., Zurawska, J., Duchowski, A. T., & Wichary, S. (2020). Pupillary and Microsaccadic Responses to Cognitive Effort and Emotional Arousal During Complex Decision Making. *Journal of Eye Movement Research*, *13*(5). https://doi.org/10.16910/JEMR.13.5.2

Kumar, A., & Nayar, K. R. (2020). COVID 19 and its mental health consequences. *Journal of Mental Health*, *30*(1), 1–2. https://doi.org/10.1080/09638237.2020.1757052

Lakatos, P., Gross, J., & Thut, G. (2019). A New Unifying Account of the Roles of Neuronal Entrainment. In *Current Biology* (Vol. 29, Issue 18, pp. R890–R905). Cell Press. https://doi.org/10.1016/j.cub.2019.07.075

Lakens, D. (2010). Movement synchrony and perceived entitativity. *Journal of Experimental Social Psychology*, *46*(5), 701–708. https://doi.org/10.1016/j.jesp.2010.03.015

Lang, M., Shaw, D. J., Reddish, P., Wallot, S., Mitkidis, P., & Xygalatas, D. (2016). Lost in the Rhythm: Effects of Rhythm on Subsequent Interpersonal Coordination. *Cognitive Science*, *40*(7), 1797–1815. https://doi.org/10.1111/cogs.12302

Lange-Malecki, B., & Treue, S. (2012). A flanker effect for moving visual stimuli. *Vision Research*, *62*, 134–138. https://doi.org/10.1016/j.visres.2012.03.016

Large, E. W., & Jones, M. R. (1999). The Dynamics of Attending: How People Track Time-Varying Events. In *Psychological Review* (Vol. 106, Issue 1, pp. 119–159).

Larsson, M., Richter, J., & Ravignani, A. (2019). Bipedal Steps in the Development of Rhythmic Behavior in Humans. *Music & Science*, *2*, 1–14. https://doi.org/10.1177/2059204319892617

Launay, J., Dean, R. T., & Bailes, F. (2013). Synchronization Can Influence Trust Following Virtual Interaction. *Experimental Psychology*, *60*(1), 53–63. https://doi.org/10.1027/1618-3169/a000173

Launay, J., Tarr, B., & Dunbar, R. I. M. (2016). Synchrony as an Adaptive Mechanism for Large-Scale Human Social Bonding. *Ethology*, *122*(10), 779–789. https://doi.org/10.1111/eth.12528

Le Merrer, J., Becker, J. A. J., Befort, K., & Kieffer, B. L. (2009). Reward processing by the opioid system in the brain. *Physiological Reviews*, *89*(4), 1379–1412. https://doi.org/10.1152/physrev.00005.2009

Leder, H. (2013). Next steps in neuroaesthetics: Which processes and processing stages to study? *Psychology of Aesthetics, Creativity, and the Arts*, *7*(1), 27–37. https://doi.org/10.1037/a0031585

Leder, H., Bär, S., & Topolinski, S. (2012). Covert Painting Simulations Influence Aesthetic Appreciation of Artworks. *Psychological Science*, *23*(12), 1479–1481. https://doi.org/10.1177/0956797612452866

Lee, H., Launay, J., & Stewart, L. (2020). Signals through music and dance: Perceived social bonds and formidability on collective movement. *Acta Psychologica*, *208*(June), 103093. https://doi.org/10.1016/j.actpsy.2020.103093

Lenggenhager, B., & Lopez, C. (2015). Vestibular Contributions to the Sense of Body, Self, and Others. In *Open Mind* (Vol. 23, Issue January, pp. 1–38). https://doi.org/10.15502/9783958570023

Lick, D. J., & Johnson, K. L. (2013). Fluency of visual processing explains prejudiced evaluations following categorization of concealable identities. *Journal of Experimental Social Psychology*, *49*(3), 419–425. https://doi.org/10.1016/j.jesp.2013.01.001

Lick, D. J., & Johnson, K. L. (2015). The Interpersonal Consequences of Processing Ease: Fluency as a Metacognitive Foundation for Prejudice. *Current Directions in Psychological Science*, *24*(2), 143–148. https://doi.org/10.1177/0963721414558116

Lick, D. J., & Johnson, K. L. (2016). Perceptually Mediated Preferences and Prejudices. *Psychological Inquiry*, *27*(4), 335–340. https://doi.org/10.1080/1047840X.2016.1215211

Lumsden, J., Miles, L. K., & Macrae, C. N. (2014). Sync or sink? Interpersonal synchrony impacts self-esteem. *Frontiers in Psychology*, *5*(SEP), 1–11. https://doi.org/10.3389/fpsyg.2014.01064

Macpherson, M. C., Fay, N., & Miles, L. K. (2020). Seeing synchrony: A replication of the effects of task-irrelevant social information on perceptions of interpersonal coordination. *Acta Psychologica*, *209*(June), 103140. https://doi.org/10.1016/j.actpsy.2020.103140

Macpherson, M. C., Marie, D., Schön, S., & Miles, L. K. (2020). Evaluating the interplay between subclinical levels of mental health symptoms and coordination dynamics. *British Journal of Psychology*, *111*(4), 782–804. https://doi.org/10.1111/bjop.12426

Macrae, C. N., Duffy, O. K., Miles, L. K., & Lawrence, J. (2008). A case of hand waving: Action synchrony and person perception. *Cognition*, *109*(1), 152–156. https://doi.org/10.1016/j.cognition.2008.07.007

Macrae, C. N., Hewstone, M., & Griffiths, R. J. (1993). Processing load and memory for stereotype-based information. *European Journal of Social Psychology*, *23*(May 1992), 77–87.

Maddigan, M. E., Sullivan, K. M., Halperin, I., Basset, F. A., & Behm, D. G. (2019). High tempo music prolongs high intensity exercise. *PeerJ*, *2019*(1). https://doi.org/10.7717/peerj.6164

Makin, A. D. J., Pecchinenda, A., & Bertamini, M. (2012). Implicit affective evaluation of visual symmetry. *Emotion*, *12*(5), 1021–1030. https://doi.org/10.1037/a0026924

Malmierca, M. S., Sanchez-Vives, M. V., Escera, C., & Bendixen, A. (2014). Neuronal adaptation, novelty detection and regularity encoding in audition. *Frontiers in Systems Neuroscience*, *8*(JUNE), 1–9. https://doi.org/10.3389/fnsys.2014.00111

Mathôt, S. (2018). Pupillometry: Psychology, Physiology, and Function. *Journal of Cognition*, *1*(1), 1–23. https://doi.org/10.5334/joc.18

Mathôt, S., Fabius, J., Van Heusden, E., & Van der Stigchel, S. (2018). Safe and sensible preprocessing and baseline correction of pupil-size data. *Behavior Research Methods*, *50*(1), 94–106. https://doi.org/10.3758/s13428-017-1007-2

McEllin, L., Knoblich, G., & Sebanz, N. (2020). Synchronicities that shape the perception of joint action. *Scientific Reports*, *10*(1), 1–10. https://doi.org/10.1038/s41598-020-72729-6

McFee, B., Metsai, A., McVicar, M., Balke, S., Thome, C., Raffel, C., Zalkow, F., Malek, A., Dana, Lee, K., Nieto, O., Ellis, D., Mason, J., Battenberg, E., Seyfarth, S., Yamamoto, R., Viktorandreevichmorozov, Choi, K., Moore, J., & Thassilo. (2022). *Librosa* (0.9.1). Zenodo.

McKinney, M. F., & Moelants, D. (2006). Ambiguity in tempo perception: what draws listeners to different metrical levels? *Music Perception*, *24*(2), 155–166.

McNeill, W. H. (1997). *Keeping Together in Time: Dance and Drill in Human History*. Harvard University Press.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, *27*(1), 415–444. https://doi.org/10.1146/annurev.soc.27.1.415

Meghanathan, R. N., van Leeuwen, C., & Nikolaev, A. R. (2015). Fixation duration surpasses pupil size as a measure of memory load in free viewing. *Frontiers in Human Neuroscience*, *8*(JAN), 1063. https://doi.org/10.3389/FNHUM.2014.01063/BIBTEX

Mehr, S. A., Krasnow, M. M., Bryant, G. A., & Hagan. (2021). Origins of music in credible signaling. *Behavioral and Brain Sciences*, 1–30.

Mehr, S. A., Singh, M., Knox, D., Ketter, D. M., Pickens-Jones, D., Atwood, S., Lucas, C., Jacoby, N., Egner, A. A., Hopkins, E. J., Howard, R. M., Hartshorne, J. K., Jennings, M. v., Simson, J., Bainbridge, C. M., Pinker, S., O'Donnell, T. J., Krasnow, M. M., & Glowacki, L. (2019). Universality and diversity in human song. *Science*, *366*(6468), eaax0868. https://doi.org/10.1126/science.aax0868

Merker, B. H. (1999). Synchronous chorusing and the origins of music. *Musicae Scientiae*, *3*(1_suppl), 59–73. https://doi.org/10.1177/10298649000030s105

Merker, B. H., Madison, G. S., & Eckerdal, P. (2009). On the role and origin of isochrony in human rhythmic entrainment. *Cortex*, *45*(1), 4–17. https://doi.org/10.1016/j.cortex.2008.06.011

Micheyl, C., Hunter, C., & Oxenham, A. J. (2010). Auditory stream segregation and the perception of across-frequency synchrony. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(4), 1029–1039. https://doi.org/10.1037/a0017601

Miles, L. K., Griffiths, J. L., Richardson, M. J., & Macrae, C. N. (2010). Too late to coordinate: contextual influences on behavioral synchrony. *European Journal of Social Psychology*, *40*, 52–60. https://doi.org/10.1002/ejsp.721

Miles, L. K., Nind, L. K., Henderson, Z., & Macrae, C. N. (2010). Moving memories: Behavioral synchrony and memory for self and others. *Journal of Experimental Social Psychology*, *46*(2), 457–460. https://doi.org/10.1016/j.jesp.2009.12.006

Miles, L. K., Nind, L. K., & Macrae, C. N. (2009). The rhythm of rapport: Interpersonal synchrony and social perception. *Journal of Experimental Social Psychology*, *45*(3), 585–589. https://doi.org/10.1016/j.jesp.2009.02.002

Mills, P. F., Van Der Steen, M. C., Schultz, B. G., & Keller, P. E. (2015). Individual Differences in Temporal Anticipation and Adaptation during Sensorimotor Synchronization. *Timing and Time Perception*, *3*(1–2), 13–31. https://doi.org/10.1163/22134468-03002040

Mogan, R., Fischer, R., & Bulbulia, J. A. (2017). To be in synchrony or not? A meta-analysis of synchrony's effects on behavior, perception, cognition and affect. *Journal of Experimental Social Psychology*, *72*(October), 13–20. https://doi.org/10.1016/j.jesp.2017.03.009

Monin, B. (2003). The Warm Glow Heuristic: When Liking Leads to Familiarity. *Journal of Personality and Social Psychology*, *85*(6), 1035–1048. https://doi.org/10.1037/0022-3514.85.6.1035

Montagu, J. (2017). How Music and Instruments Began: A Brief Overview of the Origin and Entire Development of Music, from Its Earliest Stages. *Frontiers in Sociology*, *2*. https://doi.org/10.3389/fsoc.2017.00008

Morgan, P., Macken, B., Toet, A., Bompas, A., Bray, M., Rushton, S., & Jones, D. (2020). Distraction for the eye and ear. *Theoretical Issues in Ergonomics Science*, *21*(6). https://doi.org/10.1080/1463922X.2020.1712493

Mü Llensiefen, D., Gingras, B., Musil, J., & Stewart, L. (2014). The Musicality of Non-Musicians: An Index for Assessing Musical Sophistication in the General Population. *PLoS ONE*, *9*(2), 89642. https://doi.org/10.1371/journal.pone.0089642

Myers, D. G. (2007). *Psychology* (8th ed.). Worth Publishers.

Nadal, M., Munar, E., Marty, G., & Cela-Conde, C. (2010). Visual complexity and beauty appreciation: Explaining the divergence of results. *Empirical Studies of the Arts*, *28*(2), 173–191. https://doi.org/10.2190/EM.28.2.d

Nietzsche, F. (1908). *Menschliches, Allzumenschliches: Ein Buch für freie Geister*. https://en.wikisource.org/wiki/Human_All-Too-Human

Nilsson, J., & Thorstensson, A. (1987). Adaptability in frequency and amplitude of leg movements during human locomotion a t different speeds. *Acta Physiologica Scandinaciva*, *129*, 107–114.

North, A. C., & Hargreaves, D. J. (1995). Subjective complexity, familiarity, and liking for popular music. *Psychomusicology: A Journal of Research in Music Cognition*, *14*(1–2), 77–93. https://doi.org/10.1037/h0094090

North, A. C., Hargreaves, D. J., & McKendrick, J. (1999). The Influence of In-Store Music on Wine Selections. *Journal of Applied Psychology*, *84*(2), 271–276.

Nothdurft, H. C. (2002). Attention shifts to salient targets. *Vision Research*, *42*(10), 1287–1306. https://doi.org/10.1016/S0042-6989(02)00016-0

Nowak, M. A. (2006). Five Rules for the Evolution of Cooperation. *Science*, *314*(5805), 1560–1563. https://doi.org/10.1126/science.1133755

Nunes, J. C., Ordanini, A., & Valsesia, F. (2015). The power of repetition: repetitive lyrics in a song increase processing fluency and drive market success. *Journal of Consumer Psychology*, *25*(2), 187–199. https://doi.org/10.1016/j.jcps.2014.12.004

Obhi, S. S., & Sebanz, N. (2011). Moving together: Toward understanding the mechanisms of joint action. *Experimental Brain Research*, *211*(3–4), 329–336. https://doi.org/10.1007/s00221-011-2721-0

Okasha, S. (2006). *Evolution and the Levels of Selection*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199267972.001.0001

Onderdijk, K. E., Swarbrick, D., van Kerrebroeck, B., Mantei, M., Vuoskoski, J. K., Maes, P. J., & Leman, M. (2021). Livestream Experiments: The Role of COVID-19, Agency, Presence, and Social Context in Facilitating Social Connectedness. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.647929

Overy, K., & Molnar-Szakacs, I. (2009). Being Together in Time: Musical Experience and the Mirror Neuron System. *Music Perception*, *26*(5), 489–504. https://doi.org/10.1525/mp.2009.26.5.489

Palacios, E. R., Isomura, T., Parr, T., & Friston, K. J. (2019). The emergence of synchrony in networks of mutually inferring neurons. *Scientific Reports*, 1–14. https://doi.org/10.1038/s41598-019-42821-7

Pearce, E., Launay, J., & Dunbar, R. I. M. (2015). The ice-breaker effect: Singing mediates fast social bonding. *Royal Society Open Science*, *2*(10). https://doi.org/10.1098/rsos.150221

Pearce, E., Wlodarski, R., Machin, A., & Dunbar, R. I. M. (2017). Variation in the β-endorphin, oxytocin, and dopamine receptor genes is associated with different dimensions of human sociality. *Proceedings of the National Academy of Sciences*, *114*(24), E4898–E4898. https://doi.org/10.1073/pnas.1708178114

Pecenka, N., Engel, A., & Keller, P. E. (2013). Neural correlates of auditory temporal predictions during sensorimotor synchronization. *Frontiers in Human Neuroscience*, *JUL*. https://doi.org/10.3389/fnhum.2013.00380

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/S13428-018-01193-Y/FIGURES/3

Petersch, B., & Dierkes, K. (2021). Gaze-angle dependency of pupil-size measurements in head-mounted eye tracking. *Behavior Research Methods*. https://doi.org/10.3758/s13428-021-01657-8

Pinet, S., Zielinski, C., Mathôt, S., Dufau, S., Alario, F. X., & Longcamp, M. (2017). Measuring sequences of keystrokes with jsPsych: Reliability of response times and interkeystroke intervals. *Behavior Research Methods*, *49*(3), 1163–1176. https://doi.org/10.3758/s13428-016-0776-3

Polak, R., Jacoby, N., Fischinger, T., Goldberg, D., Holzapfel, A., & London, J. (2018). Rhythmic prototypes across cultures: A comparative study of tapping synchronisation. *Music Perception*, *36*(1), 1–23. https://doi.org/10.1525/MP.2018.36.1.1

Pool, E., Sennwald, V., Delplanque, S., Brosch, T., & Sander, D. (2016). Measuring wanting and liking from animals to humans: A systematic review. *Neuroscience and Biobehavioral Reviews*, *63*, 124–142. https://doi.org/10.1016/j.neubiorev.2016.01.006

Pressnitzer, D., Suied, C., & Shamma, S. A. (2011). Auditory scene analysis: the sweet music of ambiguity. *Frontiers in Human Neuroscience*, *5*(December), 1–11. https://doi.org/10.3389/fnhum.2011.00158

Pronin, E., & Wegner, D. M. (2006). Manic Thinking: Independent effects of thought speed and thought content on mood. *Psychological Science*, *17*(9), 807–813. https://doi.org/10.1111/j.1467-9280.2006.01786.x

Purves, D., Paydarfar, J. A., & Andrews, T. J. (1996). The wagon wheel illusion in movies and reality. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(8), 3693–3697. https://doi.org/10.1073/pnas.93.8.3693

R Core Team. (2022). *R: a language and environment for statistical computing*. https://www.R-project.org/

Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality*, *41*(1), 203–212. https://doi.org/10.1016/j.jrp.2006.02.001

Ramstead, M. J. D., Veissière, S. P. L., & Kirmayer, L. J. (2016). Cultural affordances: Scaffolding local worlds through shared intentionality and regimes of attention. *Frontiers in Psychology*, *7*(JUL), 1–21. https://doi.org/10.3389/fpsyg.2016.01090

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*(1), 79–87. https://doi.org/10.1038/4580

Ravignani, A., Bowling, D., & Fitch, W. T. (2014). Chorusing, synchrony and the evolutionary functions of rhythm. *Frontiers in Psychology*, *5*(SEP), 1–15. https://doi.org/10.3389/fpsyg.2014.01118

Ravreby, I., Shilat, Y., & Yeshurun, Y. (2022). Liking as a balance between synchronization, complexity and novelty. *Scientific Reports*, *12*(1), 3181. https://doi.org/10.1038/s41598-022-06610-z

Reber, R., & Schwarz, N. (1999). Effects of Perceptual Fluency on Judgments of Truth. *Consciousness and Cognition*, *8*(3), 338–342. https://doi.org/10.1006/ccog.1999.0386

Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: is beauty in the perceiver's processing experience? *Personality and Social Psychology Review*, *8*(4), 364–382.

Reber, R., & Unkelbach, C. (2010). The Epistemic Status of Processing Fluency as Source for Judgments of Truth. *Review of Philosophy and Psychology*, *1*(4), 563–581. https://doi.org/10.1007/s13164-010-0039-7

Reber, R., Winkielman, P., & Schwarz, N. (1998). Effects of Perceptual Fluency on Affective Judgments. *Psychological Science*, *9*(1), 45–48. https://doi.org/10.1111/1467-9280.00008

Reddish, P., Fischer, R., & Bulbulia, J. (2013). Let's Dance Together: Synchrony, Shared Intentionality and Cooperation. *PLoS ONE*, *8*(8). https://doi.org/10.1371/journal.pone.0071182

Reilly, J., Kelly, A., Kim, S. H., Jett, S., & Zuckerman, B. (2019). The human task-evoked pupillary response function is linear: Implications for baseline response scaling in pupillometry. *Behavior Research Methods*, *51*(2), 865–878. https://doi.org/10.3758/s13428-018-1134-4

Rennung, M., & Göritz, A. S. (2016). Prosocial Consequences of Interpersonal Synchrony. *Zeitschrift Für Psychologie*, *224*(3), 168–189. https://doi.org/10.1027/2151-2604/a000252

Repp, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin and Review*, *12*(6), 969–992. https://doi.org/10.3758/BF03206433

Richmond, J. H. (2019). *Drumming up harmony: how rhythmic synchronisation fosters pro-sociality* [Doctoral]. University of Melbourne.

Rideaux, R., Badcock, D. R., Johnston, A., & Edwards, M. (2016). Temporal synchrony is an effective cue for grouping and segmentation in the absence of form cues. *Journal of Vision*, *16*(11), 23. https://doi.org/10.1167/16.11.23

Robertson, C., Tarr, B., Kempnich, M., & Dunbar, R. I. M. (2017). Rapid partner switching may facilitate increased broadcast group size in dance compared with conversation groups. *Ethology*, *123*(10), 736–747. https://doi.org/10.1111/eth.12643

Samuels, E., & Szabadi, E. (2008). Functional Neuroanatomy of the Noradrenergic Locus Coeruleus: Its Roles in the Regulation of Arousal and Autonomic Function Part I: Principles of Functional Organisation. *Current Neuropharmacology*, *6*(3), 235–253. https://doi.org/10.2174/157015908785777229

Savage, P. E., Brown, S., Sakai, E., & Currie, T. E. (2015). Statistical universals reveal the structures and functions of human music. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(29), 8987–8992. https://doi.org/10.1073/pnas.1414495112

Savage, P. E., Loui, P., Tarr, B., Schachner, A., Glowacki, L., Mithen, S., & Fitch, W. T. (2021). Music as a coevolved system for social bonding. *Behavioral and Brain Sciences*, *44*, 1–22. https://doi.org/10.1017/S0140525X20000333

Schäfer, K., Saarikallio, S., & Eerola, T. (2020). Music May Reduce Loneliness and Act as Social Surrogate for a Friend: Evidence from an Experimental Listening Study. *Music & Science*, *3*, 205920432093570. https://doi.org/10.1177/2059204320935709

Scheflen, A. E. (1964). The Significance of Posture in Communication Systems. *Psychiatry*, *27*(4), 316–331. https://doi.org/10.1080/00332747.1964.11023403

Schmeck, A., Opfermann, M., van Gog, T., Paas, F., & Leutner, D. (2015). Measuring cognitive load with subjective rating scales during problem solving: differences between immediate and delayed ratings. *Instructional Science*, *43*, 93–114. https://doi.org/10.1007/s11251-014-9328-3

Schneider, D. M., Sundararajan, J., & Mooney, R. (2018). A cortical filter that learns to suppress the acoustic consequences of movement. *Nature*, *561*(7723), 391–395. https://doi.org/10.1038/s41586-018-0520-5

Schulter, M., & Burmeister, K. (2015). *Pythagorean Tuning and Medieval Polyphony*. https://www.researchgate.net/publication/282294166_Pythagorean_Tuning_and_Medieval_Polyphony_Author_Margo_Schulter_10_June_1998_edited_by_Karl_Burmeister_contribution_primarily_graphics

Schwarz, N. (2004). Metacognitive Experiences in Consumer Judgment and Decision Making. *Journal of Consumer Psychology*, *14*(4), 332–348.

Schwarz, N. (2012). Feelings-as-Information Theory. In *Handbook of Theories of Social Psychology* (Issue January, pp. 289–308). SAGE Publications Ltd. https://doi.org/10.4135/9781446249215.n15

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: Bodies and minds moving together. *Trends in Cognitive Sciences*, *10*(2), 70–76. https://doi.org/10.1016/j.tics.2005.12.009

Shamay-Tsoory, S. G., & Abu-Akel, A. (2016). The Social Salience Hypothesis of Oxytocin. *Biological Psychiatry*, *79*(3), 194–202. https://doi.org/10.1016/j.biopsych.2015.07.020

Shamay-Tsoory, S. G., Saporta, N., Marton-Alper, I. Z., & Gvirts, H. Z. (2019). Herding Brains: A Core Neural Mechanism for Social Alignment. *Trends in Cognitive Sciences*, *23*(3), 174–186. https://doi.org/10.1016/j.tics.2019.01.002

Shamma, S. A., Elhilali, M., & Micheyl, C. (2011). Temporal coherence and attention in auditory scene analysis. *Trends in Neurosciences*, *34*(3), 114–123. https://doi.org/10.1016/j.tins.2010.11.002

Shih, Y. N., Huang, R. H., & Chiang, H. Y. (2012). Background music: Effects on attention performance. *Work*, *42*(4), 573–578. https://doi.org/10.3233/WOR-2012-1410

Silva, R. R., Garcia-Marques, T., & Reber, R. (2017). The informative value of type of repetition: Perceptual and conceptual fluency influences on judgments of truth. *Consciousness and Cognition*, *51*, 53–67. https://doi.org/10.1016/j.concog.2017.02.016

Soundy, A., Bojner Horwitz, E., Korošec, K., & Theorell, T. (2022). *Can Dance and Music Make the Transition to a Sustainable Society More Feasible?* https://doi.org/10.3390/bs12010011

Street, N., Forsythe, A. M., Reilly, R., Taylor, R., & Helmy, M. S. (2016). A Complex Story: Universal Preference vs. Individual Differences Shaping Aesthetic Response to Fractals Patterns. *Frontiers in Human Neuroscience*, *10*, 1–14. https://doi.org/10.3389/fnhum.2016.00213

Stupacher, J. (2019). The experience of flow during sensorimotor synchronization to musical rhythms. *Musicae Scientiae*, *23*(3), 348–361. https://doi.org/10.1177/1029864919836720

Stupacher, J., Maes, P. J., Witte, M., & Wood, G. (2017). Music strengthens prosocial effects of interpersonal synchronization – If you move in time with the beat. *Journal of Experimental Social Psychology*, *72*(February), 39–44. https://doi.org/10.1016/j.jesp.2017.04.007

Sullivan, P. J., & Blacker, M. (2017). The effect of different phases of synchrony on pain threshold in a drumming task. *Frontiers in Psychology*, *8*(JUN), 1–7. https://doi.org/10.3389/fpsyg.2017.01034

Sullivan, P. J., Rickers, K., & Gammage, K. L. (2014). The effect of different phases of synchrony on pain threshold. *Group Dynamics*, *18*(2), 122–128. https://doi.org/10.1037/gdn0000001

Sun, Z., & Firestone, C. (2020). The Dark Room Problem. *Trends in Cognitive Sciences*, *24*(5), 346–348. https://doi.org/10.1016/j.tics.2020.02.006

Swarbrick, D., Seibt, B., Grinspun, N., & Vuoskoski, J. K. (2021). Corona Concerts: The Effect of Virtual Concert Characteristics on Social Connection and Kama Muta. *Frontiers in Psychology*, *12*(June), 1–21. https://doi.org/10.3389/fpsyg.2021.648448

Świrski, L., & Dodgson, N. A. (2013). A fully-automatic, temporal approach to single camera, glint-free 3D eye model fitting. *Pervasive Eye Tracking and Mobile Eye-Based Interaction (PETMEI), January*.

Tarr, B., Launay, J., Cohen, E., & Dunbar, R. I. M. (2015). Synchrony and exertion during dance independently raise pain threshold and encourage social bonding. *Biology Letters*, *11*(10). https://doi.org/10.1098/rsbl.2015.0767

Tarr, B., Launay, J., & Dunbar, R. I. M. (2014). Music and social bonding: 'Self-other' merging and neurohormonal mechanisms. *Frontiers in Psychology*, *5*(SEP), 1–10. https://doi.org/10.3389/fpsyg.2014.01096

Tarr, B., Launay, J., & Dunbar, R. I. M. (2016). Silent disco: dancing in synchrony leads to elevated pain thresholds and social closeness. *Evolution and Human Behavior*, *37*(5), 343–349. https://doi.org/10.1016/j.evolhumbehav.2016.02.004

Tarr, B., Slater, M., & Cohen, E. (2018). Synchrony and social connection in immersive Virtual Reality. *Scientific Reports*, *8*(1), 1–8. https://doi.org/10.1038/s41598-018-21765-4

Theeuwes, J., Kramer, A. F., & Belopolsky, A. v. (2004). Attentional set interacts with perceptual load in visual search. *Psychonomic Bulletin & Review*, *11*(4), 697–702. https://link.springer.com/content/pdf/10.3758/BF03196622.pdf

Tinio, P. P. L., & Leder, H. (2009). Natural Scenes Are Indeed Preferred, but Image Quality Might Have the Last Word. *Psychology of Aesthetics, Creativity, and the Arts*, *3*(1), 52–56. https://doi.org/10.1037/a0014835

Topolinski, S. (2010). Moving the eye of the beholder: Motor components in vision determine aesthetic preference. *Psychological Science*, *21*(9), 1220–1224. https://doi.org/10.1177/0956797610378308

Topolinski, S. (2013). The sources of fluency. In C. Unkelbach & R. Greifeneder (Eds.), *The experience of thinking: How the fluency of mental processes influences cognition and behaviour* (pp. 33–49). Psychology Press.

Topolinski, S., & Reber, R. (2010). Immediate truth - Temporal contiguity between a cognitive problem and its solution determines experienced veracity of the solution. *Cognition*, *114*(1), 117–122. https://doi.org/10.1016/j.cognition.2009.09.009

Topolinski, S., & Strack, F. (2009). Motormouth: Mere exposure depends on stimulus-specific motor simulations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(2), 423–433. https://doi.org/10.1037/a0014504

Treisman, A. M. (1977). Focused attention in the perception and retrieval of multidimensional stimuli. In *Perception & Psychophysics* (Vol. 22, Issue 1). https://link.springer.com/content/pdf/10.3758/BF03206074.pdf

Treisman, A. M., & Gelade, G. (1980). A Feature-Integration Theory of Attention. In *COGNITIVE PSYCHOLOGY* (Vol. 12).

Tunçgenç, B., & Cohen, E. (2016). Movement synchrony forges social bonds across group divides. *Frontiers in Psychology*, *7*(MAY), 1–12. https://doi.org/10.3389/fpsyg.2016.00782

Tunçgenç, B., & Cohen, E. (2018). Interpersonal movement synchrony facilitates pro-social behavior in children's peer-play. *Developmental Science*, *21*(1), 1–9. https://doi.org/10.1111/desc.12505

Tunçgenç, B., Cohen, E., & Fawcett, C. (2015). Rock With Me: The Role of Movement Synchrony in Infants' Social and Nonsocial Choices. *Child Development*, *86*(3), 976–984. https://doi.org/10.1111/cdev.12354

UK Department for Digital, Culture, M. and S. (2018). A connected society: A strategy for tackling loneliness. In *HM Government*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/750909/6.4882_DCMS_Loneliness_Strategy_web_Update.pdf

Valdesolo, P., & DeSteno, D. (2011). Synchrony and the social tuning of compassion. *Emotion*, *11*(2), 262–266. https://doi.org/10.1037/a0021302

Vallat, R. (2018). Pingouin: statistics in Python. *Journal of Open Source Software*, *3*(31), 1026. https://doi.org/10.21105/joss.01026

Van de Cruys, S. (2017). Affective Value in the Predictive Mind. *Open MIND*, 1–21. https://doi.org/10.15502/9783958573253

Van de Cruys, S., & Wagemans, J. (2011). Putting Reward in Art: A Tentative Prediction Error Account of Visual Art. *I-Perception*, *2*(9), 1035–1062. https://doi.org/10.1068/i0466aap

van der Steen, M. C. (Marieke), & Keller, P. E. (2013). The ADaptation and Anticipation Model (ADAM) of sensorimotor synchronization. *Frontiers in Human Neuroscience*, *7*, 253. https://doi.org/10.3389/fnhum.2013.00253

Van Dyck, E., Vansteenkiste, P., Lenoir, M., Lesaffre, M., & Leman, M. (2014). Recognizing induced emotions of happiness and sadness from dance movement. *PLoS ONE*, *9*(2), 1–7. https://doi.org/10.1371/journal.pone.0089773

van Noorden, L. P. A. S. (1975). *Temporal coherence in the perception of tone sequences*. Institute for Perceptual Research.

Varlet, M., Williams, R., & Keller, P. E. (2020). Effects of pitch and tempo of auditory rhythms on spontaneous movement entrainment and stabilisation. *Psychological Research*, *84*, 568–584. https://doi.org/10.1007/s00426-018-1074-8

Vasil, J., Badcock, P. B., Constant, A., Friston, K. J., & Ramstead, M. J. D. (2020a). A World unto Itself: Human Communication as Active Inference. *Frontiers in Psychology*, *11*(March), 1–26. https://doi.org/10.3389/fpsyg.2020.00417

Vasil, J., Badcock, P. B., Constant, A., Friston, K., & Ramstead, M. J. D. (2020b). A World Unto Itself: Human Communication as Active Inference. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.00417

Veissiere, S. (2021). Deritualization in the digital age. *SPA2021: Interrogating Inequalities*. https://spa.americananthro.org/past-meetings/2021-spa-biennial-2/

Veissière, S. P. L., Constant, A., Ramstead, M. J. D., Friston, K. J., & Kirmayer, L. J. (2019). Thinking Through Other Minds: A Variational Approach to Cognition and Culture. *Behavioral and Brain Sciences*, 1–97. https://doi.org/10.1017/S0140525X19001213

Vesper, C., van der Wel, R. P. R. D., Knoblich, G., & Sebanz, N. (2011). Making oneself predictable: reduced temporal variability facilitates joint action coordination. *Experimental Brain Research*, *211*(3–4), 517–530. https://doi.org/10.1007/s00221-011-2706-z

Vicaria, I. M., & Dickens, L. (2016). Meta-Analyses of the Intra- and Interpersonal Outcomes of Interpersonal Coordination. *Journal of Nonverbal Behavior*, *40*(4), 335–361. https://doi.org/10.1007/s10919-016-0238-8

Vicary, S., Sperling, M., Von Zimmermann, J., Richardson, D. C., & Orgs, G. (2017). Joint action aesthetics. *PLoS ONE*, *12*(7), 1–21. https://doi.org/10.1371/journal.pone.0180101

Vuoskoski, J. K., & Reynolds, D. (2019). Music, rowing, and the aesthetics of rhythm. *The Senses and Society*, *14*(1), 1–14. https://doi.org/10.1080/17458927.2018.1525201

Wang, Y., Guo, Y., Wang, J., Liu, Z., & Li, X. (2021). Pupillary response to moving stimuli of different speeds. *Journal of Eye Movement Research*, *14*(1), 3. https://doi.org/https://10.16910/jemr.14.1.3

Waskom, M. (2021). Seaborn: Statistical Data Visualization. *Journal of Open Source Software*, *6*(60), 3021. https://doi.org/10.21105/joss.03021

Wasley, D., Konvalinka, I., Noy, L., Levit-Binun, N., & Golland, Y. (2015). *Being in the zone: physiological markers of togetherness in joint improvisation*. https://doi.org/10.3389/fnhum.2015.00187

Weinstein, D., Launay, J., Pearce, E., Dunbar, R. I. M., & Stewart, L. (2016). Singing and social bonding: Changes in connectivity and pain threshold as a function of group size. *Evolution and Human Behavior*, *37*(2), 152–158. https://doi.org/10.1016/j.evolhumbehav.2015.10.002

Weiss, C., Herwig, A., & Schü Tz-Bosbach, S. (2011). The Self in Social Interactions: Sensory Attenuation of Auditory Action Effects Is Stronger in Interactions with Others. *PLoS ONE*, *6*(7), 22723. https://doi.org/10.1371/journal.pone.0022723

Wen, O. X., Bamford, J. S., Burger, B., Zayas, V., & Toiviainen, P. (2017). Music-induced synchrony and social affiliation in tango. *25th Anniversary Conference for the European Society for the Cognitive Sciences of Music*.

Wertheimer, M. (1938). Laws of Organisation in Perceptual Forms. In W. Ellis (Ed.), *A source book of Gestalt psychology* (pp. 71–88). Routledge & Kegan Paul. http://www.sapili.org/subir-depois/en/ps000200.pdf

Wheatley, T., & Sievers, B. (2016). Toward a neuroscience of social resonance. In J. D. Greene, I. Morrison, & M. E. P. Seligman (Eds.), *Positive Neuroscience* (pp. 37–54). Oxford University Press.

Whittlesea, B. W. A., & Leboe, J. P. (2003). Two fluency heuristics (and how to tell them apart). *Journal of Memory and Language*, *49*(1), 62–79. https://doi.org/10.1016/S0749-596X(03)00009-3

Willems, S., & van der Linden, M. (2006). Mere exposure effect: A consequence of direct and indirect fluency-preference links. *Consciousness and Cognition*, *15*(2), 323–341. https://doi.org/10.1016/j.concog.2005.06.008

Wilson, E. O. (2012). *The Social Conquest of Earth*. Liveright.

Wiltermuth, S. S., & Heath, C. (2009). Synchrony and cooperation. *Psychological Science*, *20*(1), 1–5. https://doi.org/10.1111/j.1467-9280.2008.02253.x

Winkielman, P., & Cacioppo, J. T. (2001). Mind at ease puts a smile on the face: Psychophysiological evidence that processing facilitation elicits positive affect. *Journal of Personality and Social Psychology*, *81*(6), 989–1000. https://doi.org/10.1037/0022-3514.81.6.989

Winkielman, P., Schwarz, N., Fazendeiro, T. A., & Reber, R. (2003). The hedonic marking of processing fluency: Implications for evaluative judgement. In *The Psychology of Evaluation: Affective Processes in Cognition and Emotion*.

Witek, M. A. G. (2022). The Mind Is a DJ: Rhythmic Entrainment in Beatmatching and Embodied Temporal Processing. *The Oxford Handbook of Time in Music*, *February*, 234–252. https://doi.org/10.1093/oxfordhb/9780190947279.013.13

Witek, M. A. G., Clarke, E. F., Wallentin, M., Kringelbach, M. L., & Vuust, P. (2014). Syncopation, body-movement and pleasure in groove music. *PLoS ONE*, *9*(4). https://doi.org/10.1371/journal.pone.0094446

Wlodarczyk, A., Zumeta, L., Pizarro, J. J., Bouchat, P., Hatibovic, F., Basabe, N., & Rimé, B. (2020). Perceived Emotional Synchrony in Collective Gatherings: Validation of a Short Scale and Proposition of an Integrative Measure. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.01721

Wolf, W., Launay, J., & Dunbar, R. I. M. (2016). *Joint attention, shared goals and social bonding*. *107*(295663), 322–337. https://doi.org/10.1111/bjop.12144.Joint

Woolhouse, M. H., Tidhar, D., & Cross, I. (2016). Effects on Inter-Personal Memory of Dancing in Time with Others. *Frontiers in Psychology*, *7*(February), 1–8. https://doi.org/10.3389/fpsyg.2016.00167

Zacharias, G. L., & Young, L. R. (1981). Influence of combined visual and vestibular cues on human perception and control of horizontal rotation. *Experimental Brain Research*, *41*(2), 159–171. https://doi.org/10.1007/BF00236605

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, *9*(2, Pt.2), 1–27. https://doi.org/10.1037/h0025848

Zeamer, C., & Fox Tree, J. E. (2013). The process of auditory distraction: Disrupted attention and impaired recall in a simulated lecture environment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(5), 1463–1472. https://doi.org/10.1037/a0032190

Zekveld, A. A., Koelewijn, T., & Kramer, S. E. (2018). The Pupil Dilation Response to Auditory Stimuli: Current State of Knowledge. *Trends in Hearing*, *22*, 233121651877717. https://doi.org/10.1177/2331216518777174

Zelechowska, A., Sanchez, V. G., & Jensenius, A. R. (2020a). Standstill to the 'beat': Differences in involuntary movement responses to simple and complex rhythms. *ACM International Conference Proceeding Series*, 107–113. https://doi.org/10.1145/3411109.3411139

Zelechowska, A., Sanchez, V. G., & Jensenius, A. R. (2020b). Standstill to the 'beat'. *Proceedings of the 15th International Conference on Audio Mostly*, 107–113. https://doi.org/10.1145/3411109.3411139

Zumeta, L. N., Oriol, X., Telletxea, S., Amutio, A., & Basabe, N. (2016). Collective efficacy in sports and physical activities: Perceived emotional synchrony and shared flow. *Frontiers in Psychology*, *6*(JAN), 1–12. https://doi.org/10.3389/fpsyg.2015.01960

6 Appendices

Appendices for all chapters, where applicable, may be found below.

6.1 Appendices for Chapter 2

6.1.1 Participant information and consent form

Welcome to this study on motion perception. Before we proceed, please note that this study cannot be completed on a mobile device. Please ensure you are on a laptop or desktop computer in a quiet environment. Close any other tabs or background applications, and put your phone on silent.

This study should take approximately 10 minutes to complete. You will first be given some more detailed information about the study, and then asked some standard demographic questions. This will be followed by a perception task, requiring a keyboard response to various moving images.

The next page may look long and scary, but please read all the way through and answer the questions at the bottom. Press the arrow button to continue.

... but before we continue, please do check that you are on a laptop or desktop computer, with a keyboard, and that there are no background applications, phones, televisions, housemates, or other distractions that may grab your attention during the task. Press the arrow button to continue.

Before you begin the study, we will need you to consent to your data being included in our research. Please read through the following information sheet.

General Information

The aim of this research study is to understand fundamental processes in the perception of synchronised movement. It is the first phase of a larger study which will investigate how

people are able to coordinate action. In this study, we (Joshua Bamford and Emma Cohen at the Institute of Cognitive and Evolutionary Anthropology, University of Oxford) are investigating how people process moving images.

We appreciate your interest in participating in this research. You have been invited to participate as you are over the age of 18 and have normal or corrected vision (e.g., with glasses or lenses). Please read through these terms before agreeing to participate. You may ask any questions before taking part by contacting the researcher (details below).

In this online study, you will answer some basic questions about yourself, and then be asked to complete a short perception task in which you will have to categorise some moving images. It should take around 10 minutes. No background knowledge is required, full instructions are given prior to the task, and you will have an opportunity to practice and familiarise yourself with the set-up first. We will be measuring the speed and accuracy at which you can categorise the images. From this we hope to learn about how people perceive movement.

The results of this study may be published in the public domain, and anonymous data collected may be made publicly available, in the interests of making science more open and accountable.

Do I have to take part?

Please note that your participation is voluntary. You may withdraw at any point during the questionnaire for any reason by closing the browser before you complete the questionnaire.

How will your data be used?

Your answers will be completely anonymous, and we will use all reasonable endeavours to keep them confidential.

Your data will be stored in a password-protected file and may be used in academic publications. We have included a 'Prefer not to say' option for each set of questions if you prefer not to answer a particular question. Research data will be stored for three years after publication or public release.

The data that we collect from you may be transferred to, and stored or processed at, a destination outside the European Economic Area ('EEA'). By submitting your personal data, you agree to this transfer, storing or processing.

Who will have access to your data?

Below is some more general information about how your data will be used.

The University of Oxford is the data controller with respect to your personal data, and as such will determine how your personal data is used in the study. The University will process your personal data for the purpose of the research. Research is a task that we perform in the public interest. Further information about your rights with respect to your personal data is available here.

We would like your permission to use your anonymised data in future studies, and to share data with other researchers (e.g. in online databases). Any personal information that could identify you will be removed or changed before files are shared with other researchers or results are made public.

Responsible members of the University of Oxford and funders may be given access to data for monitoring and/or audit of the study to ensure we are complying with guidelines, or as otherwise required by law.

This questionnaire is for an DPhil project. The principal researcher is Joshua Bamford, who is at the Institute of Cognitive and Evolutionary Anthropology at the University of Oxford. This project is being completed under the primary supervision of Dr Emma Cohen.

This project has been reviewed by, and received ethics clearance through, the University of Oxford Central University Research Ethics Committee (reference number: SAME_C1A_19_002).

What if there is a problem?

If you have a concern about any aspect of this project, please speak to the researcher (joshua.bamford@anthro.ox.ac.uk) or their supervisor (emma.cohen@anthro.ox.ac.uk), who will do their best to answer your query. The researcher should acknowledge your concern within 10 working days and give you an indication of how they intend to deal with it. If you remain unhappy or wish to make a formal complaint, please contact the relevant Chair of the Research Ethics Committee at the University of Oxford:

Chair, School of Anthropology and Museum Ethnography Departmental Research Ethics Committee; Email: hod@anthro.ox.ac.uk; Address: SAME DREC, University of Oxford, 51-53 Banbury Road, Oxford OX2 6PE

The Chair will seek to resolve the matter in a reasonably expeditious manner.

Please note that you may only participate in this survey if you are 18 years of age or over and have normal or corrected vision (eg. with glasses or contacts).

I confirm that I am 18 years of age or over.

- ✓ Yes

I confirm that I have normal or corrected-to-normal vision.

- ✓ Yes

I confirm that I am definitely using a desktop or laptop computer, and have closed any background applications that may distract me while completing the task.

- ✓ Yes

I agree to take part in the study.

- ✓ Yes

6.1.2 Demographic questions

How old are you (in years)?

What is your country of origin?

What is your country of residence?

What is your gender?

- o Female
- o Male
- o Other
- o Prefer not to say

Are you left or right handed?

- o Left
- o Right

What is your highest completed level of education?

- o High School
- o Some university
- o Bachelor degree
- o Master's degree

- Doctorate

6.2 Appendices for Chapter 3

6.2.1 Advertisement

Title: Shifting Attention in Motion Perception

Duration: 60 Minutes

Credits: 1 Credit (undergraduate pool only)

Brief Description: Help us understand how action coordination may influence attention in this finger tapping, eye-tracking study.

Description:

You are invited to participate in a study to help us understand how people allocate attention while coordinating actions with others. If you choose to participate, you will be asked to complete a synchronous tapping task, in which you will drum on the desk along to a video. At the same time, you will be given a standard visual attention task, while your eye movements are being tracked. There will be 15 minutes for briefing and calibration of the eye-tracker. The task itself should take around 30 minutes, to be followed by a short questionnaire and then debriefing.

Eligibility requirements:

- Individuals must be over 18 to participate.
- Individuals cannot participate if they have any injuries that impair their ability to move their arm/arms.
- Individuals must have normal or corrected-to-normal vision and hearing.
- Individuals cannot participate if they require glasses to view a computer screen.

Must be over 18; must have normal or corrected-to-normal vision and hearing; must be able to use a computer without glasses (contact lenses are OK); cannot have injuries that impair ability to move arms.

Please arrive well rested, and refrain from intense activity immediately before the study.

COVID-19 requirements:

- Please be advised that upon arrival you will be asked a set of screening questions
- If you have returned from overseas or been in contact with a confirmed case of COVID-19 in the last 14-days, you will NOT be able to participate in the research
- If you are feeling unwell, have a fever, or had a temperature in the past 14-days, you will NOT be able to participate in the research

- If you are being tested for COVID-19, or have come into contact with someone who is under-going COVID-19 testing, you will NOT be able to participate in the research until the result have come back negative
- You are encouraged to install the COVIDSafe app and have it running while attending the study session, however this is not a requirement
- Please limit attendance to just yourself
- Please observe social distancing, and appropriate hand hygiene and cough etiquette

6.2.2 Participant information sheet

Project title: Shifting Attention in Movement Perception
Name of Researchers: Joshua Bamford (PhD student), Emma Cohen (external supervisor), Lynden Miles

Invitation:

You are invited to take part in this project because you are over the age of 18, have normal or corrected vision and hearing, and are not wearing glasses. Please talk to the researchers if you have a vision or hearing impairment which may make it difficult for you to complete the perception tasks in this study. The study will take place in the Sanders Building at UWA, and is a collaboration between UWA and the University of Oxford, UK.

Aim of the Study (What is the project about?)

People all over the world engage in coordinated action, when working, dancing or playing together. This aim of this study is to understand fundamental processes in the perception of synchronised movement. It forms part of a larger study which will investigate the social effects of coordinated action.

What does participation involve?

If you agree to take part, you will be asked to answer a few questions about yourself, and then complete a tapping task. In total, it should take not more than an hour. The task will involve watching a video of two people tapping on the desk in front of them. You will be asked to tap along with one of these people. At the same time, you will also complete a visual attention task. We are not interested in how well you perform on the task, but in how the different stimuli interact with each other. Just try your best to complete the task as quickly and accurately as possible, while tapping in time with your target. Your responses to the visual attention task will be recorded by the computer, and your tapping will be audio recorded by the contact microphone in front of you. While performing the task, your eye movements will also be video recorded by an eye tracking headset. The whole study will be broken up into multiple rounds, and there will be a short break at the end of each round. You will have the opportunity to ask the researcher any questions you may have before you begin. If you are still happy to take part, you will then be asked to sign a consent form.

Voluntary Participation and Withdrawal from the Study

Your participation in this study is entirely voluntary. You may withdraw at any time during or after the study, without giving an explanation and with no consequences. Any data collected from you would be destroyed if you decide to withdraw.

Your privacy

The information you provide as part of the study is the research data. Any research data from which you can be identified (e.g. your name, date of birth), is known as personal data. It does not include data where the identity has been removed (anonymous data).

We minimise our use of personal data in the study as much as possible and any identifiable responses will be anonymised. Note that we will be recording your eyes in this study, however this is only to measure your pupil size and gaze direction – no video that could be used to identify you from your pupils is recorded.

The research data will be stored confidentially on a local hard disc, and then backed up to labarchives.com. Personal data, will be stored confidentially on a local hard disc, and separately from any research data, while consent forms will be kept in a locked cabinet.

The researcher and supervisor will have access to personal and research data. Responsible members of the University of Oxford or the University of Western Australia may be given access to data for monitoring and/or audit of the research.

All research data will be stored for at 7 years after publication or public release of the work of the research. We may retain and store your personal data for an additional period of time as necessary for the purposes of the study, and for further research.

Your personal data may be transferred to, and stored at, a destination outside Australia or the European Economic Area. We will make sure that any data transfer is done securely and with a similar level of data protection as required under UK and Australian law.

We would like your permission to use anonymised data in future studies, and to share data with other researchers (e.g. in online databases). All personal information that could identify you will be removed or changed before information is shared with other researchers or results are made public.

The University of Oxford is committed to the dissemination of its research for the benefit of society and the economy and, in support of this commitment, has established an online archive of research materials. This archive includes digital copies of student theses successfully submitted as part of a University of Oxford postgraduate degree programme. Holding the archive online gives easy access for researchers to the full text of freely available theses, thereby increasing the likely impact and use of that research.

The research will be written up as a thesis. On successful submission of the thesis, it will be deposited both in print and online in the University archives, to facilitate its use in future research. The thesis will be openly accessible. Findings from this study may also be published in academic journals.

Possible Benefits

This study is investigating the mechanisms involved in how we coordinate action with others, and may contribute towards our knowledge of fundamental human social behaviour. There will be no direct benefit to you from taking part in this research. However, we are happy to share our findings with you once the research is completed, if you are interested. Please

indicate if you wish for us to save your contact details so that we may share the findings with you. Note that contact details will never be stored with the data collected during the study.

Possible Risks and Risk Management Plan

There are minimal risks in taking part in this study. It is possible that you will feel fatigued by the end of the study. In order to reduce this, you will be given breaks between rounds, and we encourage you to sit in a comfortable position while completing the tasks. You will be given a separate information sheet about the possible risks of COVID-19.

Reimbursement (include this section only if relevant)

You will receive credit points for your participation in this study.

Contacts

If you would like to participate, to discuss any aspect of this study, or to express any concerns about this study, please feel free to contact Joshua Bamford (Joshua.bamford@anthro.ox.ac.uk), or external supervisor (emma.cohen@anthro.ox.ac.uk).

6.2.3 Participant consent form

I, _____ have read the information provided and any questions I have asked have been answered to my satisfaction. I agree to participate in this research project, realising that I may withdraw at any time without reason and without prejudice.

I understand that all identifiable information that I provide is treated as confidential and will not be released by the investigator in any form that may identify me unless I have consented to this. The only exception to this principle of confidentiality is if this information is required by law to be released.

I understand that my drumming, but not my voice, will be audio-recorded. Yes ☐ No ☐

I understand that my pupils, but not my whole face, will be measured with a camera, but that no identifiable information will be stored. Yes ☐ No ☐

I understand that the above recordings are a requirement of participation in this study. Yes ☐ No ☐

_____ _____

Participant signature Date

Approval to conduct this research has been provided by the University of Western Australia, in accordance with its ethics review and approval procedures. Any person considering participation in this research project, or agreeing to participate, may raise any questions or issues with the researchers at any time.

In addition, any person not satisfied with the response of researchers may raise ethics issues or concerns, and may make any complaints about this research project by contacting the Human Ethics Office at the University of Western Australia on (08) 6488 3703 or by emailing to humanethics@uwa.edu.au

All research participants are entitled to retain a copy of any Participant Information Form and/or Participant Consent Form relating to this research project.

6.2.4 Instructions for participants

Welcome to this study. Please read the following instructions carefully. Press any key to continue.

In this study, you will see a video of two hands tapping on a desk. With your left hand, tap along with the hand on the left. The tempo may change sometimes, but try your best to stay in time with the hand on the left.

While tapping along to the video with your left hand, you will be also be shown some letters on the screen. Amongst the other letters, there will be either an 'N' or an 'M'. If you see an 'N', press the 'N' key. If you see an 'M', press the 'M' key. You should respond with your right hand. There is no time limit, but try to answer as quickly as possible.

You will be given six sessions, with a break in between each of them. Before starting, you'll get three practice sessions. Remember: Tap along to the video on the left of the screen, using your left hand. Respond using the 'N' or 'M' key with your right hand.

That is the end of the practice trials. Please take the time now to ask the experimenter if you have any questions. When you are ready, we will continue to the real sessions. Press 'space' to continue.

That is the end of this task. Please call the experimenter for further instructions.

6.2.5 Debrief sheet

Shifting attention in movement perception - Debrief

The purpose of this study was to measure the effect of synchronous tapping on visual attention.

How was this tested?

You may have noticed that the two people in the background video were tapping in time with each other for some trials, but not in others. While tapping along with the target person on your left, you were also presented with a standard visual search task. This is what is known as a dual-task paradigm, in which you perform two tasks concurrently, and the effect of one task is measured with the other task. It assumes that we have limited cognitive resources, and that the two tasks must compete for those resources. In this case, we were looking at the effect of trying to keep tapping in time with the target person in the video would have on the visual search task. In addition, you were having your eye-movements tracked. This allows us to see

where your visual attention is directed, but also how hard you were concentrating on the task, as our pupils change size when we are more focused.

Hypotheses

There were three conditions in this study: one hand tapping, both hands tapping in synchrony, both hands tapping out of synchrony. We expect that performance on the visual search task would be worsened in the non-synchrony condition, as the movement of the non-synchronous target would be a distraction, making the tapping task more difficult. We expect no difference between the other two conditions, since in the synchronous tapping condition both hands are moving together and should provide no distraction. With the eye-tracking, we expect that your pupils would be changing size more rapidly during the non-synchronous condition than in the other two, reflecting the greater degree of concentration required.

Why is this important to study?

Synchronised movement is an important feature of human social interaction and of human cultures throughout history. All cultures engage in sustained, synchronised action in the form of singing, dancing and drumming together. It is thought that these activities help to create and sustain social bonds within a group, but the mechanisms are poorly understood. This research aims to test a theory that we like to be in synchrony with others because it is easier to process their actions. This is important, as it may help us to understand a fundamental aspect of human social activity, and could have applications in how we design music therapy interventions or community-based music projects.

What if I want to know more?

If you are interested in learning more about this research, or would like to see a summary of our findings when it is complete, please contact Joshua Bamford (joshua.bamford@anthro.ox.ac.uk).

If you feel discomfort regarding any of the material or would like to talk to someone regarding personal issues raised by participating in this study, you can contact the student welfare team at 6488 2424. Alternative support can be reached at HeadSpace (1800 650 890) which provides support for individuals between 12-25 years old.

Finally, please remember that your participation in this study is entirely voluntary and you can withdraw from the study at any time without giving a reason and without loss of course credit or reimbursement. If you wish to withdraw your data following this debrief, please let one of the research team know.

Approval to conduct this research has been provided by the University of Western Australia, in accordance with its ethics review and approval procedures. Any person considering participation in this research project, or agreeing to participate, may raise any questions or issues with the researchers at any time. In addition, any person not satisfied with the response of researchers may raise ethics issues or concerns, and may make any complaints about this research project by contacting the Human Ethics office at UWA on (08) 6488 4703 or by emailing to humanethics@uwa.edu.au. All research participants are entitled to retain a copy of any Participant Information Form and/or Participant Consent Form relating to this research project.

This study has also been approved by the University of Oxford. If you have a concern about any aspect of this study, please speak to either Joshua Bamford (joshua.bamford@anthro.ox.ac.uk) or his supervisor (emma.cohen@anthro.ox.ac.uk), who will do their best to answer your query. The researcher should acknowledge your concern within 10 working days and give you an indication of how they intend to deal with it. If you remain unhappy or wish to make a formal complaint, please contact the relevant chair of the Research Ethics Committee at the University of Oxford who will seek to resolve the matter in a reasonably expeditious manner:
Chair, School of Anthropology and Museum Ethnography Departmental Research Ethics Committee; Email: hod@anthro.ox.ac.uk; Address: SAME DREC, University of Oxford, 51-53 Banbury Road, Oxford OX2 6PE

6.2.6 Big Five Index – 10 Item version (BFI-10)

I see myself as someone who… (5-point Likert scale)

… is reserved

… is generally trusting

… tends to be lazy

… is relaxed, handles stress well

… has few artistic interests

… is outgoing, sociable

… tends to find fault with others

… does a thorough job

… gets nervous easily

… has an active imagination

6.2.7 Goldsmiths Musical Sophistication Inventory (Gold-MSI)

Please select the most appropriate category for each statement:

(7-point Likert scale)

I am able to judge whether someone is a good singer or not

I usually know when I'm hearing a song for the first time

I find it difficult to spot mistakes in a performance of a song even if I know the tune

I can compare and discuss differences between two performances or versions of the same piece of music

I have trouble recognising a familiar song when played in a different way or by a different performer

I have never been complimented for my talents as a musical performer

I can tell when people sing or play out of time with the beat

I can tell when people sing or play out of tune

When I sing, I have no idea whether I'm in tune or not

When I hear a piece of music I can usually identify its genre

I would not consider myself a musician

I engaged in regular, daily practice of a musical instrument (including voice) for __ years.

- 0
- 1
- 2
- 3
- 4-5
- 6-9
- 10 or more

At the peak of my interest, I practised __ hours per day on my primary instrument.

- 0
- 0.5
- 1
- 1.5
- 2
- 3-4
- 5 or more

I had formal training in music theory for __ years.

- 0
- 0.5
- 1
- 2
- 3
- 4-6
- or more

I have had __ years of formal training on a musical instrument (including voice) during my lifetime.

- 0
- 0.5
- 1
- 2
- 3-5
- 6-9
- 10 or more

I can play __ musical instruments.

- 0
- 1
- 2
- 3
- 4
- 5
- 6

6.2.8 Demographic questions

How old are you (in years)?

Nationality:

Country of residence:

What is your gender?

- Male
- Female
- Other
- Prefer not to say

Are you left or right handed?

- Left
- Right

What is the highest educational qualification you have attained?

- Did not complete any school qualification
- Secondary school certificate
- Undergraduate degree or professional qualification
- Postgraduate degree
- I am still in education

If you are still in education, what is the highest qualification you expect to obtain?

- Second school certificate
- Undergraduate degree or professional qualification
- Postgraduate degree
- Not applicable

6.2.9 Supplementary Analysis

To test for presentation order effects, an LMEM was constructed including condition and trial and predicting performance (see Table *6.1*). Note that the conditions are comparative with the control condition. No difference was found between synchrony and control. However, performance was significantly lower in non-synchrony, compared with control. There was no effect of trial number or any interaction between trial number and the synchrony conditions.

Table 6.1 *Linear Mixed-Effects Model of the interaction effects between condition (synchrony vs control, non-synchrony vs control) and trial number on performance.*

Predictors	Estimates	CI	*p*
(Intercept)	0.72	0.67 – 0.77	<0.001
Non-synchrony	-0.1	-0.16 – -0.04	0.001
Synchrony	0.02	-0.04 – 0.08	0.452
Trial	0	-0.00 – 0.01	0.191
Non-synchrony * Trial	0.01	-0.00 – 0.02	0.113
Condition Synchrony * Trial	0	-0.01 – 0.01	0.796
Random Effects			
σ2	0		
τ00 Participant	0.01		

ICC	0.82
N Participant	72
Observations	393
Marginal R^2 / Conditional R^2	0.053 / 0.828

6.2.10 Video stimuli

The three video stimuli for this task can be found in Table *6.2* below.

Table 6.2 *List of video stimuli used in Study 2.*

Condition	Link
Synchrony	https://osf.io/g9y5n
Non-synchrony	https://osf.io/gfza2
Control	https://osf.io/3gh6s

6.3 Appendices for Chapter 4

6.3.1 Advertisement

Title: Attention Control in a Drumming Task

Duration: 30 Minutes

Credits: 0.5 Credit (undergraduate pool only)

Brief Description: Help us understand how action coordination may influence attention in this drumming and eye-tracking study.

Description:

You are invited to participate in a study to help us understand how people allocate attention while coordinating actions with others. If you choose to participate, you will be asked to complete a task in which you will drum on the desk along to a video. Afterwards, you will be asked a few questions about the experience. The study should take around 30 minutes.

Eligibility requirements:

- Individuals must be over 18 to participate.
- Individuals cannot participate if they have any injuries that impair their ability to move their arm/arms.
- Individuals must have normal or corrected-to-normal vision and hearing.

- Individuals cannot participate if they require glasses to view a computer screen.

Must be over 18; must have normal or corrected-to-normal vision and hearing; must be able to use a computer without glasses (contact lenses are OK); cannot have injuries that impair ability to move arms.

Please arrive well rested, and refrain from intense activity immediately before the study.

COVID-19 requirements:

- Please be advised that upon arrival you will be asked a set of screening questions
- If you have returned from overseas or been in contact with a confirmed case of COVID-19 in the last 14-days, you will NOT be able to participate in the research
- If you are feeling unwell, have a fever, or had a temperature in the past 14-days, you will NOT be able to participate in the research
- If you are being tested for COVID-19, or have come into contact with someone who is under-going COVID-19 testing, you will NOT be able to participate in the research until the result have come back negative
- You are encouraged to install the COVIDSafe app and have it running while attending the study session, however this is not a requirement
- Please limit attendance to just yourself
- Please observe social distancing, and appropriate hand hygiene and cough etiquette

6.3.2 Participant information sheet

Project title: Attention Control in a Drumming Task
Name of Researchers: Joshua Bamford (PhD student), Emma Cohen (external supervisor), Lynden Miles

Invitation:

You are invited to take part in this project because you are over the age of 18, have normal or corrected vision and hearing, and are not wearing glasses. Please talk to the researchers if you have a vision or hearing impairment which may make it difficult for you to complete the perception tasks in this study. The study will take place in the Sanders Building at UWA, and is a collaboration between UWA and the University of Oxford, UK.

Aim of the Study (What is the project about?)

People all over the world engage in coordinated action, when working, dancing or playing together. This aim of this study is to understand fundamental processes in the perception of synchronised movement. It forms part of a larger study which will investigate the social effects of coordinated action.

What does participation involve?

If you agree to take part, you will be asked to answer a few questions about yourself, and then complete a tapping task. In total, it should take not more than 30 minutes. The task will involve watching a video of two people tapping on a drum in front of them. You will be

asked to tap along with one of these people, using the drum provided. Your tapping will be audio recorded by a contact microphone. While performing the task, your eye movements will also be video recorded by an eye tracking headset. The whole study will be broken up into multiple rounds. After each round you will be asked some questions. Before you begin you will be asked to sign a consent form, and you will be fully debriefed afterwards.

Voluntary Participation and Withdrawal from the Study

Your participation in this study is entirely voluntary. You may withdraw at any time during or after the study, without giving an explanation and with no consequences. Any data collected from you would be destroyed if you decide to withdraw.

Your privacy

The information you provide as part of the study is the research data. Any research data from which you can be identified (e.g. your name, date of birth), is known as personal data. It does not include data where the identity has been removed (anonymous data).

We minimise our use of personal data in the study as much as possible and any identifiable responses will be anonymised. Note that we will be recording your eyes in this study, however this is only to measure your pupil size and gaze direction – no video that could be used to identify you from your pupils is recorded.

The research data will be stored confidentially on a local hard disc, and then backed up to labarchives.com. Personal data, will be stored confidentially on a local hard disc, and separately from any research data, while consent forms will be kept in a locked cabinet.

The researcher and supervisor will have access to personal and research data. Responsible members of the University of Oxford or the University of Western Australia may be given access to data for monitoring and/or audit of the research.

All research data will be stored for at 7 years after publication or public release of the work of the research. We may retain and store your personal data for an additional period of time as necessary for the purposes of the study, and for further research.

Your personal data may be transferred to, and stored at, a destination outside Australia or the European Economic Area. We will make sure that any data transfer is done securely and with a similar level of data protection as required under UK and Australian law.

We would like your permission to use anonymised data in future studies, and to share data with other researchers (e.g. in online databases). All personal information that could identify you will be removed or changed before information is shared with other researchers or results are made public.

The University of Oxford is committed to the dissemination of its research for the benefit of society and the economy and, in support of this commitment, has established an online archive of research materials. This archive includes digital copies of student theses successfully submitted as part of a University of Oxford postgraduate degree programme. Holding the archive online gives easy access for researchers to the full text of freely available theses, thereby increasing the likely impact and use of that research.

The research will be written up as a thesis. On successful submission of the thesis, it will be deposited both in print and online in the University archives, to facilitate its use in future research. The thesis will be openly accessible. Findings from this study may also be published in academic journals.

Possible Benefits

This study is investigating the mechanisms involved in how we coordinate action with others, and may contribute towards our knowledge of fundamental human social behaviour. There will be no direct benefit to you from taking part in this research. However, we are happy to share our findings with you once the research is completed, if you are interested. Please indicate if you wish for us to save your contact details so that we may share the findings with you. Note that contact details will never be stored with the data collected during the study.

Possible Risks and Risk Management Plan

There are minimal risks in taking part in this study. It is possible that you will feel fatigued by the end of the study. In order to reduce this, you will be given breaks between rounds, and we encourage you to sit in a comfortable position while completing the tasks. You will be given a separate information sheet about the possible risks of COVID-19.

Reimbursement

If you are a psychology student at UWA, you will receive credit points for your participation in this study.

Contacts

If you would like to participate, to discuss any aspect of this study, or to express any concerns about this study, please feel free to contact Joshua Bamford (Joshua.bamford@anthro.ox.ac.uk), or external supervisor (emma.cohen@anthro.ox.ac.uk).

6.3.3 Participant consent form

I, _____ have read the information provided and any questions I have asked have been answered to my satisfaction. I agree to participate in this research project, realising that I may withdraw at any time without reason and without prejudice.

I understand that all identifiable information that I provide is treated as confidential and will not be released by the investigator in any form that may identify me unless I have consented to this. The only exception to this principle of confidentiality is if this information is required by law to be released.

I understand that my drumming, but not my voice, will be audio-recorded. Yes ☐ No ☐

I understand that my pupils, but not my whole face, will be measured with a camera, but that no identifiable information will be stored. Yes ☐ No ☐

I understand that the above recordings are a requirement of participation in this study. Yes ☐ No ☐

_____ _____

Participant signature Date

Approval to conduct this research has been provided by the University of Western Australia, in accordance with its ethics review and approval procedures. Any person considering participation in this research project, or agreeing to participate, may raise any questions or issues with the researchers at any time.

In addition, any person not satisfied with the response of researchers may raise ethics issues or concerns, and may make any complaints about this research project by contacting the Human Ethics Office at the University of Western Australia on (08) 6488 3703 or by emailing to humanethics@uwa.edu.au

All research participants are entitled to retain a copy of any Participant Information Form and/or Participant Consent Form relating to this research project.

6.3.4 Instructions for participants

Welcome to this study. Please read the following instructions carefully, as you cannot move back through the slides. Press any key to continue.

In this task you will see pairs of people drumming side-by-side. For each video you will be asked to tap the drum in front of you with your dominant hand, in time with the person on the LEFT of the screen. Press any key to continue.

After each video, you will be asked some questions. Press any key to continue.

Most of these questions will be answered on a slider like the one below. In each case, please click and hold on the slider, and drag the marker to the desired location. When you release the mouse button your answer will be recorded and you will move to the next slide. Try this now to continue.

< slider />

In addition to these sliders, sometimes you will also be asked to indicate your response using circles on the screen that you can drag into position. Press any key to continue.

For these questions, you will see a circle like this on the right. When you click on or near the ‘A’ circle, a ‘B’ circle will appear. Press any key to continue.

Once the ‘B’ circle appears, keep holding down the mouse button, and drag the circle into the appropriate position. Your response is recorded when you release the mouse button. Try this on the next slide. Press any key to continue.

Click next to the ‘A’ circle and drag the ‘B’ circle into the desired position. When you release the mouse button, your answer will be recorded and you will move to the next slide.

< IOS slider />

There will be six videos, each of one minute in length. These videos are taken from a standard set of stimuli for use in research. Please take a moment to get comfortable with the drum. On the next slide you will see an example of the stimuli, except with only one drummer. Press any key to continue.

6.3.5 Post-task questions

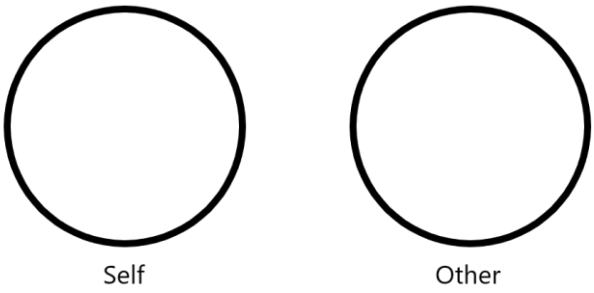
Please rate how much you like the person on the LEFT.

Not at all ---------- very much

Please rate how connected you felt to the person on the LEFT.

Not at all ---------- very connected

Please click next to the ‘other’ circle and drag the ‘self’ circle into the position that best describes your relationship to the person on the LEFT.



Please rate the difficulty of the task.

Very easy ---------- very difficult

(All questions used continuous scales from 0 to 1 in the experiment itself and were then scaled by 100 for analysis.)

6.3.6 Debrief sheet

Attention Control in a Drumming Task - Debrief

The purpose of this study was to measure the effect of synchronous drumming on visual attention, and subsequently on feelings of social affiliation.

How was this tested?

You may have noticed that the two people in the video were tapping in time with each other for some trials, but not in others. While drumming, you were having your eye-movements tracked. This allows us to see where your visual attention is directed, but also how hard you were concentrating on the task, as our pupils change size when we are more focused. In addition, there were a series of measures of social bonding and affiliation after each trial. Some participants saw videos with two different drums, while others the videos both had the same drums.

Hypotheses

There were three synchrony conditions in this study, expressed as tempo ratios: synchronised (120:120), complementary (120:80), non-synchrony (120:113). We expect that the synchrony and complementary conditions should be easier than non-synchrony. If so, we expect that participants would have more pupil activity and larger pupil diameter during the non-synchrony condition. We also expect that participants will also like the target drummer more in the synchrony condition, as expressed on the range of self-report social measures.

Why is this important to study?

Synchronised movement is an important feature of human social interaction and of human cultures throughout history. All cultures engage in sustained, synchronised action in the form of singing, dancing and drumming together. It is thought that these activities help to create and sustain social bonds within a group, but the mechanisms are poorly understood. This research aims to test a theory that we like to be in synchrony with others because it is easier to process their actions. This is important, as it may help us to understand a fundamental aspect of human social activity, and could have applications in how we design music therapy interventions or community-based music projects.

What if I want to know more?

If you are interested in learning more about this research, or would like to see a summary of our findings when it is complete, please contact Joshua Bamford (joshua.bamford@anthro.ox.ac.uk).

If you feel discomfort regarding any of the material or would like to talk to someone regarding personal issues raised by participating in this study, you can contact the student welfare team at 6488 2424. Alternative support can be reached at HeadSpace (1800 650 890) which provides support for individuals between 12-25 years old.

Finally, please remember that your participation in this study is entirely voluntary and you can withdraw from the study at any time without giving a reason and without loss of course credit or reimbursement. If you wish to withdraw your data following this debrief, please let one of the research team know.

Approval to conduct this research has been provided by the University of Western Australia, in accordance with its ethics review and approval procedures. Any person considering participation in this research project, or agreeing to participate, may raise any questions or

issues with the researchers at any time. In addition, any person not satisfied with the response of researchers may raise ethics issues or concerns, and may make any complaints about this research project by contacting the Human Ethics office at UWA on (08) 6488 4703 or by emailing to humanethics@uwa.edu.au. All research participants are entitled to retain a copy of any Participant Information Form and/or Participant Consent Form relating to this research project.

This study has also been approved by the University of Oxford. If you have a concern about any aspect of this study, please speak to either Joshua Bamford (joshua.bamford@anthro.ox.ac.uk) or his supervisor (emma.cohen@anthro.ox.ac.uk), who will do their best to answer your query. The researcher should acknowledge your concern within 10 working days and give you an indication of how they intend to deal with it. If you remain unhappy or wish to make a formal complaint, please contact the relevant chair of the Research Ethics Committee at the University of Oxford who will seek to resolve the matter in a reasonably expeditious manner:
Chair, School of Anthropology and Museum Ethnography Departmental Research Ethics Committee; Email: hod@anthro.ox.ac.uk; Address: SAME DREC, University of Oxford, 51-53 Banbury Road, Oxford OX2 6PE

6.3.7 Big Five Index – 10 Item version (BFI-10)

I see myself as someone who… (5-point Likert scale)

… is reserved

… is generally trusting

… tends to be lazy

… is relaxed, handles stress well

… has few artistic interests

… is outgoing, sociable

… tends to find fault with others

… does a thorough job

… gets nervous easily

… has an active imagination

6.3.8 Goldsmiths Musical Sophistication Inventory (Gold-MSI)

Please select the most appropriate category for each statement:

(7-point Likert scale)

I am able to judge whether someone is a good singer or not

I usually know when I'm hearing a song for the first time

I find it difficult to spot mistakes in a performance of a song even if I know the tune

I can compare and discuss differences between two performances or versions of the same piece of music

I have trouble recognising a familiar song when played in a different way or by a different performer

I have never been complimented for my talents as a musical performer

I can tell when people sing or play out of time with the beat

I can tell when people sing or play out of tune

When I sing, I have no idea whether I'm in tune or not

When I hear a piece of music I can usually identify its genre

I would not consider myself a musician

I engaged in regular, daily practice of a musical instrument (including voice) for __ years.

- 0
- 1
- 2
- 3
- 4-5
- 6-9
- 10 or more

At the peak of my interest, I practised __ hours per day on my primary instrument.

- 0
- 0.5
- 1
- 1.5
- 2
- 3-4
- 5 or more

I had formal training in music theory for __ years.

- 0
- 0.5
- 1
- 2
- 3

- 4-6
- or more

I have had __ years of formal training on a musical instrument (including voice) during my lifetime.

- 0
- 0.5
- 1
- 2
- 3-5
- 6-9
- 10 or more

I can play __ musical instruments.

- 0
- 1
- 2
- 3
- 4
- 5
- 6

6.3.9 Demographic questions

How old are you (in years)?

Nationality:

Country of residence:

What is your gender?

- Male
- Female
- Other
- Prefer not to say

Are you left or right handed?

- Left
- Right

What is the highest educational qualification you have attained?

- Did not complete any school qualification
- Secondary school certificate
- Undergraduate degree or professional qualification
- Postgraduate degree
- I am still in education

If you are still in education, what is the highest qualification you expect to obtain?

- Second school certificate
- Undergraduate degree or professional qualification
- Postgraduate degree
- Not applicable

6.3.10 Supplementary analyses

For exploratory purposes, a Linear Mixed-Effects Model (LMEM) of bonding was constructed including all tempo ratio, timbre and IOI variance variables as predictors. This initial model failed to converge. Following the procedure of Barr et al. (2013), variables were removed until a viable model was constructed. This modelled the effects of tempo ratio, difficulty, and pupil size on bonding, with participant as a random effect (Table *6.3*). A significant effect ($p < .05$) was found of the interaction between difficulty and pupil size (Figure 6.1). Thus, as ratings of difficulty increased, there was a stronger negative effect of pupil size on bonding. This suggests that the task may not have been difficult enough to register changes in pupil size that would correspond with any bonding effects in most participants. Note also that when difficulty and pupil size are included in the model there is no direct effect of tempo ratio, which is corroborated in the mediation analysis.

Table 6.3 *Linear mixed-effects model predicting Bonding from Tempo Ratio, Difficulty and Pupil Size.*

Predictors	**Estimates**	**CI**	***p***
(Intercept)	58.60	45.24 – 71.96	<.001
Tempo Ratio (Complementary)	-3.10	-11.96 – 5.75	.491
Tempo Ratio (Non-synchrony)	1.97	-7.63 – 11.57	.687
Difficulty	-0.03	-0.38 – 0.32	.872
Pupil Size	1.09	-1.82 – 4.01	.462
TR(Complementary) * Difficulty	0.03	-0.04 – 0.11	.382
TR(Non-synchrony) * Difficulty	-0.02	-0.10 – 0.05	.527
TR(Complementary) * Pupil Size	0.66	-1.35 – 2.67	.518
TR(Non-synchrony) * Pupil Size	-0.59	-2.73 – 1.55	.590
Difficulty * Pupil Size	-0.09	-0.17 – -0.01	.026
Random Effects			
σ2	100.26		
τ00 Participant	270.51		
ICC	0.73		
N Participant	82		
Observations	466		
Marginal R^2 / Conditional R^2	0.179 / 0.778		

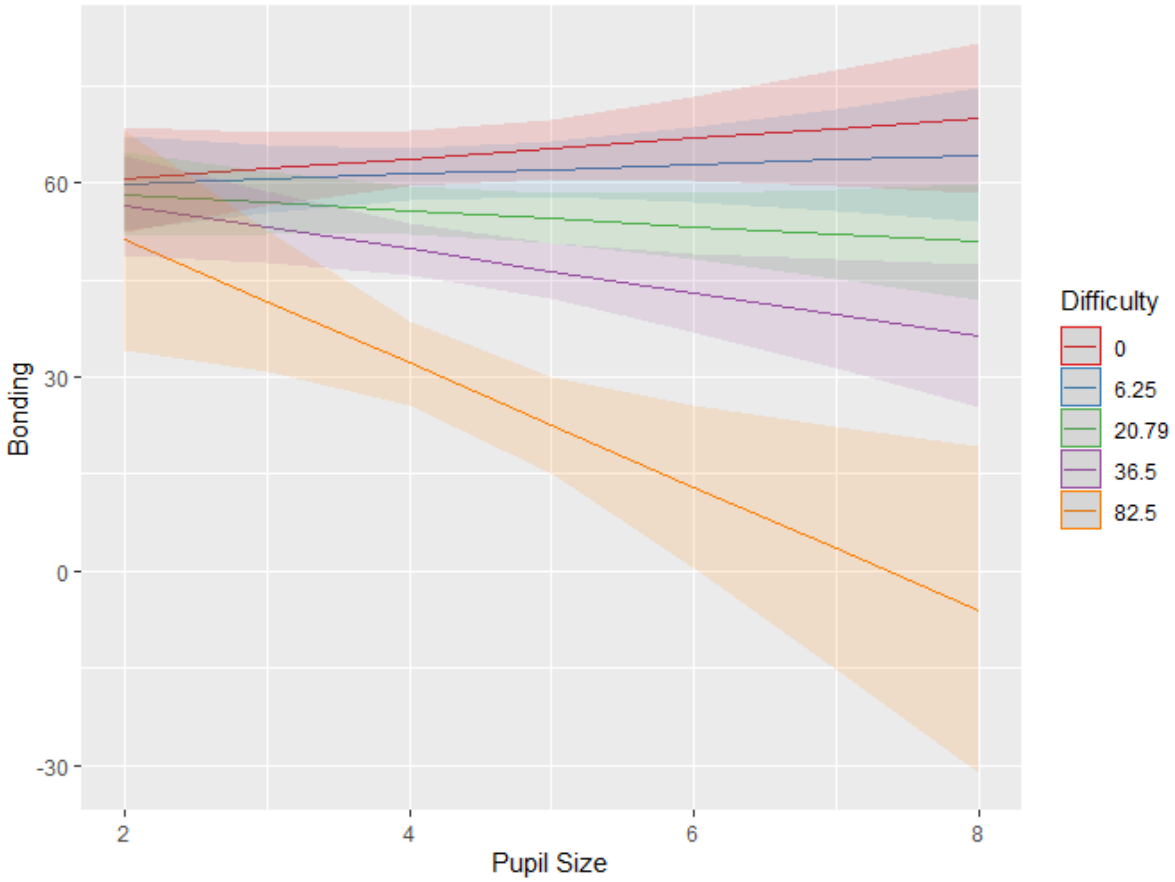


Figure 6.1 Interaction effect of pupil size and difficulty on bonding. Difficulty was divided into 5 bands, each corresponding to the colours in the legend. Shading shows the confidence interval of the model.

6.3.11 Video stimuli

Six different conditions were created between the two factors: tempo ratio (three levels) and timbre (two levels). The individual drummers that appeared in each condition was randomised. Example videos from each condition can be found in Table *6.4*.

Table 6.4 *Example videos for each condition in Study 3.*

	Synchrony	Complementary	Non-synchrony
Matched drums	https://osf.io/bn849	https://osf.io/gxsrd	https://osf.io/4rwtb
Unmatched drums	https://osf.io/m732j	https://osf.io/eth4q	https://osf.io/85hxp