# Towards a Psychology of Surveillance: Do ‘Watching Eyes’ Affect Behaviour?

**Keith P Dear**
Magdalen College
University of Oxford

A thesis submitted for the degree of
*Doctor of Philosophy*
Trinity Term 2018

# Abstract

## Towards a Psychology of Surveillance: Do ‘Watching Eyes’ Affect Behaviour?

Surveillance matters. Chapter 1 shows we are today watched more closely than at any time in modernity. Philosophers have long speculated that just *feeling* watched might change our behaviour, but empirical evidence for this is limited. This thesis extends the evidence by exploring ‘the watching eyes effect’ – a phenomenon identified by psychologists which suggests that the presence of pictures of eyes can trigger us to feel more watched and act more prosocially. The research presented also tests the effect of the ‘mere presence’ of a camera. In Chapter 1 we see that there *is* significant evidence that we are hypersensitive to gaze and feeling watched. However, a laboratory study reported in Chapter 2 finds no effect of eye cues or the presence of a camera on prosociality. In Chapter 3 a systematic review of the available watching eyes literature develops the hypothesis that *if,* as others suggest, watching eyes cue us to feel watched and worry about our reputation we might find they have more consistent effects on antisocial behaviour. A meta-analysis of 15 experiments supports this contention, finding a 35% reduction in antisocial behaviour when eye cues are present. Seeking to explore this further, Chapter 4 reports three studies finding that eye cues (or a camera) do not reduce antisocial behaviour in the laboratory. Chapter 5 reports the results of a field study finding consistent reductions in crime at 9 locations where signs are erected – three of which are statistically significant with crime falling 27%, 100%, and 100%. Chapter 6 reports a carefully designed large-scale field experiment in the London Borough of Ealing. Unexpectedly, over the twelve months after we put the signs up there was no cycle crime for 6 months preventing meaningful assessment of the signs’ effectiveness. In concluding, we consider the possibility that our laboratory studies show participants don’t feel watched in the presence of eye cues or a camera – that the watching eyes effect is the result of false positives and publication bias. However, noting Chapter 1’s evidence for our hypersensitivity to surveillance, and our findings in Hereford this investigation concludes that perhaps more consistent watching eyes effects can be found in the field than in the laboratory. In the end, having synthesised the psychological and interdisciplinary evidence for surveillance effects and added its empirical findings to the canon, this thesis advances us towards a psychology of surveillance, and points the way to further research.

# Acknowledgements

My thanks go first to Dr Kevin Dutton, for the gifts of sunrise whiskeys, bottle-moments, the Sorites paradox (also known as the paradox of the heap – no reference, I hope, to the work herein), for thesis advice and for his friendship. To my friend and our scientific conscience Professor Elaine Fox for keeping "DPhil" at the top of the whiteboard, keeping rigour in the research & for ensuring Kev & I made it through the last 3 years alive.

Special mention to Matt Ellen for his software genius, advice, and support. Likewise, to those who were so generous with their time and expertise in response to three years of my panicked WhatsApp requests for stats advice: Dr Anne-Wil Kruijt, Sam Parsons & Dr Maud Grol. Also, to the OCEAN lab for their friendship, advice, and support: Eda Tipura, Annabel Songco, Charlotte Booth, Emilia Boehm, Alex Temple-McCune, Desiree Spronck & Lauren Heathcote.

This thesis might be all my fault but the opportunity to write it would never have existed without the belief others have shown in me and their willingness to take risk in approving my research. Group Captain Paul Wilkins, the first to endorse and champion my research proposal to whom I owe a very great debt indeed. To General Mark Welsh, Chief of Staff of the US Air Force for taking the time to hear my pitch, read the proposal and offer his endorsement of it – without which the research might never have happened. To Air Vice Marshal (AVM) Edward Stringer and his successor as Assistant Chief of the Air Staff, AVM Richard Knighton, for giving me a chance to show this was possible and winning the support needed to make it so. To Group Captain Jim Beldon and Air Commodore Richard Barrow for their advice and support in helping me get through these three years whilst remaining (I hope) a useful and productive member of the RAF. To Lieutenant Colonel Nicole Nichols for supporting me through the application process while I was under her command. And to AVM Mike Hart whose support in 2010 manoeuvred me on to the MA programme at King's and whose example, support and guidance helped me believe this might be possible.

Thanks also to Sergeant Duncan Reynolds and Mr Charles Naylor of West Mercia Police, and to Mr Nick Massey at Ealing Council for their fellowship, embrace of structured crime prevention measures, advocacy of our shared idea, & patience with the demands of academia.

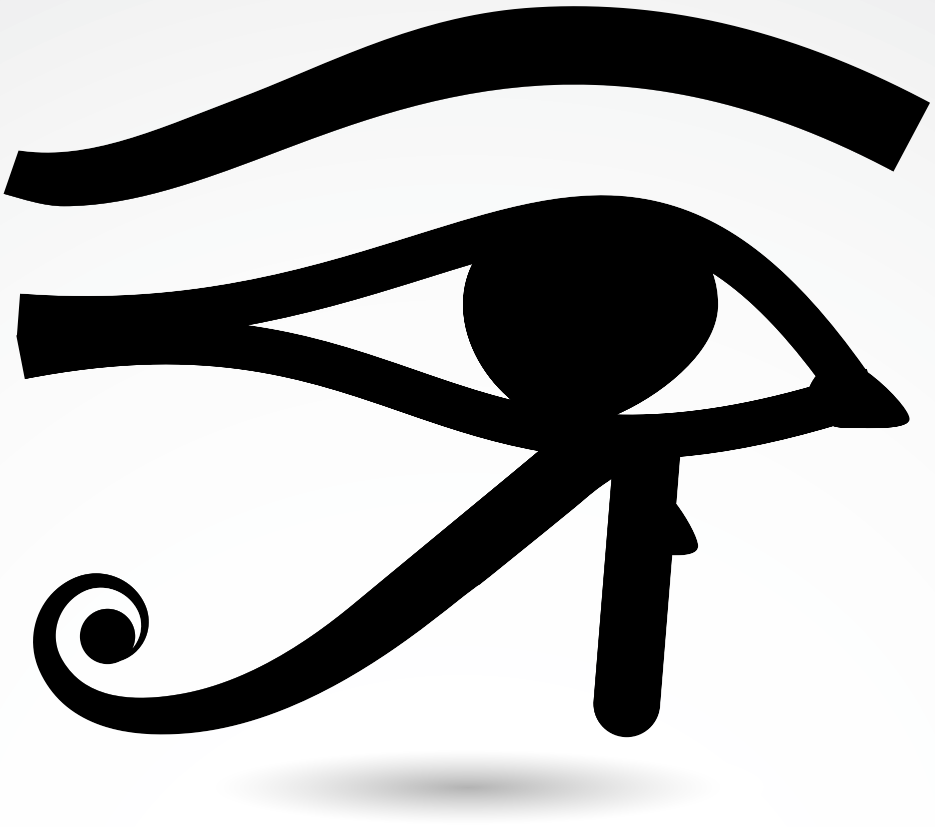
And to many others who made this research possible or helped me through it. To Cyan Collier and Peter Ward for their friendship, humour and advice. To Hugh Ward for his humanity, generosity and friendship. To Chris McLernon, Jesse Peterson & all at the USAF's European Office for Aerospace Research and Defence for their comradeship, humour, patience and advice. To other friends, old and new, on whom I rely for so much.

Last but most of all to my family. Dad – for telling me to 'look it up' more times than I can count. With a bit of luck - here's that extra 2%... Mum for her work ethic, love and belief in me. To my sister Ruth for the expert proof-reading & advice. My brother for his humour and friendship. To my Mother-in-Law Mahbub and Father-in-Law Mehdi, for their love, support and for all the free childcare!

Most of all to my beautiful, smart and strong wife Gilda: Best. Wife. Ever. & perfect Mum to our daughter, Arabella – because these stories don't mean anything…

# Declaration

This thesis comprises work conducted by Keith Patrick Dear for the degree of Doctor of Philosophy from the University of Oxford. One of the chapters, Chapter 3 contains content that is currently under review in a peer-reviewed journal, but this chapter has been edited and formatted for the thesis. Recruitment and data collection for studies 1 and 2 were carried out at Royal Air Force, Royal Navy, British Army and other Defence Establishments across the UK, as well as at civilian digital health company Incuna in Oxford. Data for studies 3 and 4 was collected through online platform Prolific Academic. Study 5 was carried out in Hereford in collaboration with West Mercia Police. Study 6 was carried out in Ealing in collaboration with Ealing Council and the London Metropolitan Police. Keith Patrick Dear has led all work included in this thesis.



Symbol of the Egyptian God Horus, or "one who is above, over". The sun was his right eye and the moon his left – always watching, day & night. It is an ancient Egyptian symbol of protection believed to originate as the symbol of the all-seeing eye of one of the earliest of Egyptian deities, Wadjet, a solar Goddess (Carabas, 2018).

# Contents

---

**1. Introduction** **1**
**1.1. Surveillance in Context** 2
1.1.1. A Philosophy of Surveillance 2
1.1.2. The Expansion of the Surveillance State 6
1.1.3. Surveillance Capitalism 14
1.1.4. Surveillance, Sousveillance, Activism & Policing 19
1.1.5. Surveillance, Peace-Keeping & Humanitarianism 24
1.1.6. So What? 27

**1.2. The Psychology of Surveillance: A Literature Review** 30
1.2.1. Surveillance Studies & Psychology 30
1.2.2. Watching Eyes Effect 32
1.2.3. Gaze in Evolutionary Perspective 33
1.2.4. Evolutionary Evidence Part 1: Ethology 34
1.2.5. Human Responsiveness to Eyes & Eye Cues 38
1.2.6. Neurophysiological Research 41
1.2.7. Evolutionary Evidence Part 2: Eyes & Evolutionary Psychology 44
1.2.8. Eye Morphology: Can we Read the Mind in the Eyes? 46
1.2.9. The Supernatural Surveillance Hypothesis 54
1.2.10 Evolutionary Evidence Part 3: Cheating in Darkness 62
1.2.11. Eye Images as Eye Cues 64
1.2.12. Mere Presence: A Psychology of Being Watched 66
1.2.13. Research Overview 69

**2. Study 1 Perceived Surveillance & Prosociality** **72**
2.1. Introduction 74
2.1.1. Public Self-Awareness, Public Self-Consciousness & Eye Cues 78
2.1.2. Personality Measure – HEXACO PI-R 79
2.2. Methods 80
2.2.1. Study Sites & Participants 80
2.2.2. Measures 80
2.2.3. Study Design 86
2.2.4. Procedure 88
2.2.5. Data Plan 89
2.3. Results 90
2.3.1. Zurich Prosocial Game – Interaction Effects Analysis 94
2.3.2. Are the Zurich Prosocial Game and Dictator Game Measuring the Same Behaviour? 99
2.3.3. Dictator Game – Interaction Effects Analysis 100
2.3.4. Snyder Self-Monitoring Scores & Prosocial Behaviour 104
2.4. Discussion 105
2.5. Limitations 110
2.6. Conclusion 112

**3. A Systematic Review & Meta-analysis of the 'Watching Eyes Effect' on Antisocial Behaviour** **114**
- 3.1. Introduction 116
  - 3.1.1. Eye Cues and Antisocial Behaviour 119
- 3.2. Methods 123
  - 3.2.1. Types of Intervention 124
  - 3.2.2. Types of outcome measure 124
  - 3.2.3. Search Strategy 124
  - 3.2.4. Meta-Analysis Study Selection 126
  - 3.2.5. Analytic Methodology & Statistical Approach 130
- 3.3. Results 135
- 3.4. Discussion 144
- 3.5. Limitations 149
- 3.6. Further Research 150
- 3.7. Conclusions 150

**4. Studies 2-4 Perceived Surveillance & Malicious Behaviour in the Joy of Destruction Game** **153**
- 4.1. Introduction 155
  - 4.1.1. Joy of Destruction Task 157
  - 4.1.2. Public Self-Consciousness 161
  - 4.1.3. Independence-Interdependence Scores 162
  - 4.1.4. Social Value Orientation 164
  - 4.1.5. Religiosity 165
  - 4.1.6. Format & Overview 166
- 4.2. Study 2. 166
  - 4.2.1. Methods 167
    - 4.2.1.1. Participants 167
    - 4.2.1.2. Design 167
    - 4.2.1.3. Materials 168
    - 4.2.1.3.1. Public Self-Consciousness Scale. 168
    - 4.2.1.3.2. Singelis Self-Construal Scale 168
    - 4.2.1.3.3. Social Value Orientation 169
    - 4.2.1.3.4. Centrality of Religion Scale 170
    - 4.2.1.4. Stimuli 170
    - 4.2.1.5. Procedure 172
    - 4.2.1.6. Data Plan 174
  - 4.2.2. Study 2 Results 175
    - 4.2.2.1. Public Self-Consciousness and Surveillance Cues 183
    - 4.2.2.2. Public Self-Consciousness as Moderator of Surveillance Cue Effects on Destruction Rates 183
    - 4.2.2.3. Independence-Interdependence Scores as Moderator of Surveillance Cue Effects on Destruction Rates 188
    - 4.2.2.4. Social Value Orientation as Moderator of Surveillance Cue Effects on Destruction Rates 186
    - 4.2.2.5. Religiosity as Moderator of Surveillance Cue Effects on Destruction Rates 188
  - 4.2.3. Study 2 Discussion 189

4.3. Study 3. 192
    4.3.1. Methods 192
        4.3.1.1. Participants 192
        4.3.1.2. Design 192
        4.3.1.3. Materials 193
        4.3.1.4. Stimuli 193
        4.3.1.5. Procedure 195
        4.3.1.6. Data Plan 195
    4.3.2. Study 3 Results 195
        4.3.2.1. Public Self-Consciousness and Surveillance Cues 199
        4.3.2.2. Public Self-Consciousness as Moderator of Surveillance Cue Effects on Destruction Rates 199
        4.3.2.3. Independence-Interdependence Scores as Moderator of Surveillance Cue Effects on Destruction Rates 201
        4.3.2.4. Social Value Orientation as Moderator of Surveillance Cue Effects on Destruction Rates 201
        4.3.2.5. Religiosity as Moderator of Surveillance Cue Effects on Destruction Rates 203
    4.3.3. Study 3 Discussion 203
4.4. Study 4. 205
    4.4.1. Methods 205
        4.4.1.1. Participants 205
        4.4.1.2. Design 205
        4.4.1.3. Materials 205
        4.4.1.4. Stimuli 205
        4.4.1.5. Procedure 206
        4.4.1.6. Data Plan 206
    4.4.2. Study 4 Results 211
        4.4.2.1. Public Self-Consciousness and Surveillance Cues 209
        4.4.2.2. Public Self-Consciousness as Moderator of Surveillance Cue Effects on Destruction Rates 209
        4.4.2.3. Independence-Interdependence Scores as Moderator of Surveillance Cue Effects on Destruction Rates 210
        4.4.2.4. Social Value Orientation as Moderator of Surveillance Cue Effects on Destruction Rates 211
        4.4.2.5. Religiosity as Moderator of Surveillance Cue Effects on Destruction Rates 212
    4.4.3. Study 4 Discussion 213
4.5. Aggregate Analysis 215
    4.5.1. Aggregate JoD Results 215
    4.5.2. Public Self-Consciousness and Surveillance Cues 222
    4.5.3. Public Self-Consciousness as Moderator of Surveillance Cue Effects on Destruction Rates 222
    4.5.4. Independence-Interdependence Scores as Moderator of Surveillance Cue Effects on Destruction Rates 224
    4.5.5. Social Value Orientation as Moderator of Surveillance Cue Effects on Destruction Rates 226
    4.5.6. Religiosity as Moderator of Surveillance Cue Effects on Destruction Rates 228
4.6. Discussion 230
4.7. Limitations 235
4.8. Further Research 239

**5. Study 5 Watching Eyes on Crime in Hereford: A Collaborative Field Experiment** **241**
5.1. Introduction 243
5.2. Method 245
    5.2.1. Design 245
    5.2.2. Study Sites 251
    5.2.3. Stimuli 247
    5.2.4. Data Recording 248
    5.2.5. Data Plan 249
5.3. Results 250
5.4. Discussion 252
5.5. Limitations 255
5.6. Further Research 256

**6. Study 6 “Cycle Thieves: We’re Watching You!” A Collaborative Field Experiment & Replication.** **258**
6.1. Introduction 260
6.2. Method 261
    6.2.1. Design 261
    6.2.2. Study Sites 261
    6.2.3. Stimuli 263
    6.2.4. Data Recording 263
    6.2.5. Data Plan 263
6.3. Results 264
6.4. Discussion 267
6.5. Limitations 268

**7. Discussion** **270**
7.1. Overview 271
7.2. Updated Meta-Analysis of the Effect of Eye Cues on Antisocial Behaviour 272
7.3. Summary of Chapters 1-6 273
7.4. Summary of Research Findings 279
7.5. Limitations 281
7.6. Further Research 283
7.7. Conclusion 285

**Bibliography** **287-321**

**Appendices** **323**
Appendix A to Chapter 1 325
Appendix B to Chapter 4 332
Appendix C to Chapter 7 337
Appendix D to Chapters 1-7 340

# Figures, Tables, Appendices

---

## Figures

*Figure 1.1. Logo of the Syrian Observatory for Human Rights*________________________22

*Figure 1.2. Stylised Eye of the Egyptian God Horus on the screen of Haley & Fessler's (2005) economic (dictator) game – where the presence of the eye cue correlated with increased participant generosity*________________________________________________33

*Figure 1.3. Cott's diagram demonstrating the inherent conspicuousness of the human eye*_35

*Figure 1.4. Crypsis: Eye Masks and Disruptive Colouration of the Eyes in Animals.*______36

*Figure 1.5. The remains of the temple at Gobekli Tepe ("Potbelly Hill") in modern South-East Turkey. Built c.9500 BCE the same time as the agricultural revolution began, or perhaps even just before, the first monuments were circles of massive T-stone pillars – megaliths built staring down across the surrounding countryside.*______________________57

*Figure 1.6. The 'wisdom eyes' of Buddha in Swayambhunath Stupa, Nepal stare down from the tower on top of the hill in all directions – with an inescapable all-seeing third eye between the brows.*________________________________________________________60

*Figure 2.1. Labelled Screenshot from the Zurich Prosocial Game (ZPG)*________________82

*Figure 2.2. Conditions: Control, Eye Cues, Camera (Present), Camera (Salient).*________87

*Figure 2.3. Mean HEXACO PI-R Emotionality Scores by Condition.*___________________92

*Figure 2.4. Mean HEXACO PI-R Agreeableness Scores by Condition.*_________________93

*Figure 2.5. Mean HEXACO PI-R Conscientiousness Scores by Condition.*_______________94

*Figure 2.6. The influence of emotionality on helping behaviours in the ZPG*_____________99

*Figure 2.7. Mean Snyder Self-Monitoring Scores by Condition.* _____________________104

*Figure 3.1. PRISMA Flow Diagram*___________________________________________126

*Figure 3.2. Forest plot of experiments included in both the limited and extended meta-analyses.*______________________________________________________________139

*Figure 3.3. Funnel plot of experiments included in both the limited and extended meta-analyses.*______________________________________________________________140

*Figure 3.4. P-curve results examining z-scores of our 15 studies for publication bias.*____141

*Figure 3.5. Funnel plot of the Fixed Effects PET-PEESE analysis for experiments 1-13.*_144

*Figure 4.1. Erasmus statue as used by Baillon et al (2013), stylised eye of Horus design used by Haley and Fessler (2005), eye posters adapted from that used in Fathi, Bateson, and Nettle (2014).* 160

*Figure 4.2. Conditions: Control, Eye Cues, Camera* 171

*Figure 4.3. Schematic of Testing Pod Used in the Study* 173

*Figure 4.4. Study 2. Distribution of Destruction Frequency by Participant* 182

*Figure 4.5. Study 2. Destruction by Mean Expected Destruction* 183

*Figure 4.6. Study 2. The influence of Public Self-Consciousness on Destruction Rates in the Joy of Destruction Game.* 188

*Figure 4.7. Study 2. The influence of Independent-Interdependent Self-Construal on Destruction Rates in the Joy of Destruction Game.* 186

*Figure 4.8. Study 2. The influence of Social Value Orientation on Destruction Rates in the Joy of Destruction Game.* 187

*Figure 4.9. Study 2. The influence of Centrality of Religion Scores on Destruction Rates in the Joy of Destruction Game.* 189

*Figure 4.10. Screenshots as seen by participants in studies 2, 3, & 4.* 194

*Figure 4.11. Study 3. The influence of Public Self-Consciousness on Destruction Rates in the Joy of Destruction Game.* 200

*Figure 4.12. Study 3. The influence of Social Value Orientation on Destruction Rates in the Joy of Destruction Game.* 202

*Figure 4.13. Study 4. The influence of Independent-Interdependent Self-Construal on Destruction Rates in the Joy of Destruction Game.* 211

*Figure 4.14. Study 4. The influence of Centrality of Religion on Destruction Rates in the Joy of Destruction Game.* 213

*Figure 4.15. Distribution (%) of Decisions to Destroy by Participant in the Joy of Destruction Task.* 221

*Figure 4.16. Aggregate Analysis: Mean Expected Destruction by Mean Destruction* 222

*Figure 4.17. Aggregate Analysis: The influence of Public Self-Consciousness on Destruction Rates in the Joy of Destruction Game.* 224

*Figure 4.18. Aggregate Analysis: The influence of Independent-Interdependent Self-Construal on Destruction Rates in the Joy of Destruction Game.* 226

*Figure 4.19. Aggregate Analysis: The influence of Social Value Orientation on Destruction Rates in the Joy of Destruction Game.* 227

*Figure 4.20. Aggregate Analysis: The influence of Centrality of Religion on Destruction Rates in the Joy of Destruction Game.* 229

*Figure 5.1. Priming Signs Used in the Hereford Study*_____________________________247

*Figure 5.2. Additional Signs Used in the Hereford Study*___________________________248

*Figure 6.1. Sign as used by Nettle et al. (2012), and the sign in location at one of the cycle sheds at Newcastle University's Campus.*_______________________________________260

*Figure 6.2. Figure 6.2. Ealing Cycle Crime Infographic*___________________________262

*Figure 6.3. Signs as used in the Ealing Study*____________________________________263

*Figure 6.4. Ealing Cycle Crime Bar Chart by 12 Monthly Period*____________________265

*Figure 6.5. Ealing Cycle Crime Bar Chart by 6 Monthly Period*_____________________266

# Tables

*Table 2.1. Summary Statistics for Study 1.* 91

*Table 2.2. Hierarchical GLM Examining Helping Behaviour in the ZPG by surveillance condition, SM Score & HEXACO Personality Index Measures.* 96-98

*Table 2.3. Hierarchical GLM Examining Dictator game donations by surveillance condition, SM Score & HEXACO Personality Index Measures.* 101-103

*Table 3.1. Studies in the Meta-Analysis exploring the effect of Eye Cues on Antisocial Behaviour (ASB).* 136

*Table 4.1. Scoring for the Centrality of Religion Scale, 1 of 2.* 170

*Table 4.2. Scoring for the Centrality of Religion Scale, 2 of 2.* 170

*Table 4.3. Joy of Destruction Study 2 Results Stratified by Condition.* 178

*Table 4.4. Study 2: Logistic Regression on the binary variable Destroyed or Not Comparing Surveillance Cue Conditions Eyes and Camera with a no surveillance cue control.* 181

*Table 4.5. Study 2: Poisson GLM Examining Destruction rates in the JoD by Condition, Public Self-Consciousness.* 184

*Table 4.6. Study 2. Poisson GLM Examining Destruction rates in the JoD by Condition, Independence-Inter-dependent Self-Construal.* 185

*Table 4.7. Study 2. Poisson GLM Examining Destruction rates in the JoD by Condition, SVO Score* 187

*Table 4.8. Study 2. Poisson GLM Examining Destruction rates in the JoD by Condition, Religiosity* 188

*Table 4.9. Summary Statistics Comparing Experiment 3 and Experiment 2.* 196

*Table 4.10. Study 3 Results Stratified by Condition* 198

*Table 4.11. Study 3. Logistic Regression: Condition vs Destroyed or Not* 199

*Table 4.12. Study 3. Poisson GLM Examining Destruction rates in the JoD by Condition, Public Self-Consciousness.* 200

*Table 4.13. Study 3. Poisson GLM Examining Destruction rates in the JoD by Condition, Independence-Interdependence Self-Construal.* 201

*Table 4.14. Study 3. Poisson GLM Examining Destruction rates in the JoD by Condition, SVO Score.* 202

*Table 4.15. Study 3. Poisson GLM Examining Destruction rates in the JoD by Condition, Religiosity.* 208

*Table 4.16. Experiment 4 Results Stratified by Condition* 207

*Table 4.17. Study 4. Logistic Regression: Condition vs Destroyed or Not* 208

*Table 4.18. Study 4. Poisson GLM Examining Destruction rates in the JoD by Condition, Public Self-Consciousness* 209

*Table 4.19. Study 4. Poisson GLM Examining Destruction rates in the JoD by Condition, Independent-Interdependent Self-Construal* 210

*Table 4.20. Study 4. Poisson GLM Examining Destruction rates in the JoD by Condition, Social Value Orientation Score* 212

*Table 4.21. Study 4. Poisson GLM Examining Destruction rates in the JoD by Condition, Religiosity* 212

*Table 4.22. Summary Statistics Comparing Studies 2, 3, & 4.* 216

*Table 4.23. JoD Results Aggregated Across All 3 Studies.* 218

*Table 4.24. Aggregate Analysis. Logistic Regression: Condition vs Destroyed or Not.* 219

*Table 4.25. Aggregate Analysis. Poisson GLM Examining Destruction rates in the JoD by Condition, Public Self-Consciousness.* 223

*Table 4.26. Aggregate Analysis. Poisson GLM Examining Destruction rates in the JoD by Condition, Independent-Interdependent Self-Construal* 225

*Table 4.27. Aggregate Analysis. Poisson GLM Examining Destruction rates in the JoD by Condition, SVO Score* 227

*Table 4.28. Aggregate Analysis. Poisson GLM Examining Destruction rates in the JoD by Condition, Religiosity* 228

*Table 5.1. Intervention Site Listings.* 246

*Table 5.2. Frequency of Antisocial Behaviour at 9 Experimental Locations.* 251

*Table 6.1. Ealing cycle theft: comparing 12 months before with 12 months after.* 264

*Table 6.2. Ealing cycle theft by 6 monthly period: First vs Third 6 Monthly Period* 266

*Table 6.3. Ealing cycle theft by 6 monthly period: Second vs Fourth 6 Monthly Period* 266

*Table C.1. Studies in the Final Meta-Analysis* 338-339

# Appendices

A. *Appendix A to Chapter 2, Study 1, Perceived Surveillance & Prosociality: Additional Analyses.* 323

B. *Appendix B to Chapter 4, Studies 2-4, Perceived Surveillance & Antisocial Behaviour: Additional Graphs and Figures.* 325

C. *Appendix C to Chapter 7. Final Meta-Analysis Table.* 337

D. *Appendix D. Note on Additional Materials for the Thesis.* 340

# Chapter 1

# 1

# Introduction

*"And Ye Shall Know the Truth, and the Truth Shall Make You Free. John VIII-XXXII"*

Central Intelligence Agency motto inscribed into the wall at the entrance to CIA Headquarters, Langley, Virginia, USA.

## 1. Introduction

## 1.1. Surveillance in Context

### 1.1.1. A Philosophy of Surveillance

The idea that surveillance can change behaviour is not new. In ancient Egypt the eye of the God Horus was worn as a talisman to ward off evil and the 'Evil Eye' has been used as a motif, hung as a charm, and worn as an amulet across cultures dating back to Classical antiquity (Bağlı & Öğüt, 2009; Dundes, 1981). The ancients seemed to believe that subtle 'eye cues' on everyday objects could act as proxies for watching gods or human gaze.

As we will see, the fields of evolutionary, cognitive and social psychology have all found some evidence to support the ideas of the ancients. We will see also how my own research suggests their ideas – that reminders of being watched can influence behaviour - cannot be dismissed and remain important as we navigate the difficult ethical and moral dilemmas that arise from the increasing ubiquity of surveillance in society and the growing capabilities of surveillance technologies in the military and security domain.

In this introduction we first frame the experimental research of Chapters 2-6 within the philosophical discussion of surveillance effects. Second, we place those philosophical concerns in the modern context and explore the wider inter-disciplinary academic study of surveillance. This framing allows us to explain the aims of the research that follows in terms of its practical relevance and application.

Perhaps the earliest secular exploration of the power of surveillance is found in Plato's 5th Century BCE account of the story of Glaucon and the ring of Gyges, a parable documenting the intrinsic centrality of observation to the formation, development, and upholding of morality (Plato, 2003, pp. 40-54). Gyges was a shepherd who descended into the underworld and returned with a ring that rendered him invisible. Quickly realizing the power the ring gave him Gyges headed off to the royal court where he promptly seduced the queen and overthrew the king. In this way, Plato suggests that hiding oneself from the gaze of society – privacy - can be a threat to social order. Plato explains that even a just man, given such a ring, would quickly begin to steal and seduce. He is clear that it is in an individual's self-interest to cheat and that only ubiquitous social surveillance enables society to function – too much privacy enables a man to cultivate a reputation for justice, without actually being just. To reap the benefits of a good, just reputation without paying the cost of acquiring one – 'seeming' to be something that one is not. Thus, it is being watched that enables societies to function cooperatively, making 'seeming' prohibitively expensive. For Plato, the all-seeing society was the ideal.

## 1. Introduction.

Philosophical and fictional discussion of state surveillance began in earnest some 200 years ago. 18th Century philosopher Jeremy Bentham's views of surveillance saw its potential utility in improving society. In his eponymous book he designed and popularised the 'Panopticon', a building where perceived continuous surveillance of the many by the few might change behaviour. Bentham's idea was that a man didn't actually need to be watched, just to feel watched "*…at every instant… and not being able to satisfy himself to the contrary, he should conceive himself to be so.*" (Bentham, 1791). Bentham pursued the idea primarily to improve the penal system, to make people work harder on menial tasks and behave better, but he also believed it had utility in hospitals, schools, sanatoriums and asylums. His ideas are enjoying a renaissance. Walmart design, glass Georgian Police Stations (Rosenburg 2013), restaurants' open kitchens, glass fronted bars, transparent government buildings, modern architectural design…are all intent on using surveillance to change behaviour (Zinnbauer, 2013).

For Sartre (1943) the mere possibility of being watched transformed one's state of mind from experiencing the self as a subject observing others to that of perceiving the self as the object of others' observation. "*I see **myself** because somebody sees me*", writes Sartre [emphasis in original text] (Sartre, 1978, pp. 259-261). According to Sartre so sensitive were we to 'the look', as he called it, that we could be tricked into seeing ourselves as object by the mere presence of a mannequin, perceiving ourselves through the mannequin's 'eyes'. The automatic, irresistible experience of perspective-taking that follows from feeling watched would transform one's world such that it was suddenly haunted by the Other's values: his fears and his judgements over which you have no control. In Sartre we find again the argument that just feeling watched changes behaviour. Though clearly valuing privacy for its ability to allow us to take a different perspective on the world, Sartre's work is one of straight description. He offers neither condemnation nor celebration of what he sees as this facet of human consciousness.

1. Introduction.

But surveillance has always had a dark side, and this is perhaps best expressed by the writer George Orwell in his dystopian conception of a future totalitarian surveillance state in the classic work, *1984*:

> *'On each landing, opposite the lift shaft, the poster with the enormous face gazed down from the wall. It was one of those pictures which are so contrived that the eyes follow you about when you move. BIG BROTHER IS WATCHING YOU, the caption beneath it ran. …the posters were plastered everywhere. …[they]…gazed down from every commanding corner. There was one on the house-front immediately opposite. BIG BROTHER IS WATCHING YOU, the caption said, while the dark eyes looked deep into Winston's own. …There was of course no way of knowing whether you were being watched at any given moment. …You have to live — did live, from habit that became instinct — in the assumption that every sound you made was overheard, and, except in darkness, every movement scrutinized'* (1949, pp. 2-5).

In *Discipline and Punish*, Foucault, like Orwell, feared that in a surveillance society conformity would be enforced less through conscious, explicit fear of detection and punishment than through the conditioning of individuals to self-censor. He repurposed Bentham's panopticon as a metaphor for modern societies. He thought that as we worried about state surveillance catching us off-guard we would internalise surveillance and suppress the expression of deviant opinions, behaviours, and eventually suppress even deviant thoughts until our conformity became unthinking, automatic. In this situation the state would have created a panoptic system of reflexive control. Citizens would conform by self-censoring but would no longer be aware they were doing so (Foucault, 1977). In Foucault's later work he came to argue more clearly that the most important and powerful means of control were not the coercive and consequential. The most powerful form of censorship was that exerted by the individual over themselves once they knew or believed they may be being watched. 'Technologies of control', Foucault suggested, were beginning to create such an environment (Manokha, 2018). However, even in his early writing he was arguing that just watching was enough:

> *'[so]… …it is not necessary to use force to constrain the convict to good behaviour, the madman to calm, the worker to work, the schoolboy to application, the patient to the observation of regulations... He who is subjected to a field of visibility, and who knows it, assumes responsibility for the constraints of power; he makes them play spontaneously upon himself; he inscribes in himself the power relation in which he simultaneously plays both roles; he becomes the principle of his own subjection'* (Foucault, 1977, pp. 202-203: cited in Manokha, 2018).

The clash of Bentham's view with that of Orwell, Foucault and modern civil liberties groups is at the heart of the discussion that follows, and underlies the empirical chapters 2-5 of this thesis. It is to these core arguments that I will return at the end. To place them in context, we next outline their contemporary relevance in the way surveillance technology is empowering authoritarian societies and the activists seeking to undermine them, how it is driving the emergence of 'surveillance capitalism', how it is changing social norms and dramatically reducing privacy in liberal democratic states, and how it empowers security agencies, militaries and peace-keeping missions.

### 1.1.2. The Expansion of the Surveillance State

One can see Bentham's positive vision of surveillance as provider of security, stability, and eradicator of free-riding in the arguments for China and Russia's surveillance systems, as well as in the discussion of the boundaries of surveillance and privacy in liberal democratic societies. In a recent article Koh Hong-Eng, Chief Scientist at the Chinese electronics company Huawei, makes the case for ubiquitous surveillance by explicitly extolling the benefits of 'Big Brother surveillance'. He argues that increasingly ubiquitous surveillance will ensure we are watched more and more closely in the future, and that we will be safer, wealthier and happier for it. Yet even Koh Hong-Eng admits that “no-one likes the idea of being watched” (Hong-Eng, 2018).

China’s social credit system is eerily reminiscent of Orwell’s surveillance state in 1984 and its aims are defended with the arguments raised by Bentham. Although the system is still being tested and varies in different administrative areas, its components are comparable. Broadly, citizens start with 300 points, and earn more by paying debts on time, making responsible purchases such as nappies or a hole-puncher, donating blood, taking public transport or reporting those who break the rules. Above 500 you get benefits – cheaper broadband, fast-track visas. Some places even promote your dating profile! However, you can also lose points - either by your own behaviour, or by associating with those with low scores, by frequenting places of ill-repute, or by having a friend who criticizes the Communist party. In Xinxiang province they even add DNA, blood samples and iris scans to the data collected on you. Xinxiang also equips police with ‘Google glass’ headsets incorporating facial recognition cameras to pick out suspects and check everyone’s background (Moody, 2018).

All this ‘watching’ enables what the Chinese Government calls “sincerity construction”, a reputational ledger owned by the state. Lose enough points, or commit particular crimes – such as jaywalking - and your face and identity number are on 5m tall TV screens in the city centre. Moreover, you can’t get away as you’re banned from trains and planes - even buses.

As if to illustrate just how far-reaching and unprecedented surveillance has become, this year, 2018, it was reported that Chinese businesses, the military and hospitals are monitoring the brainwaves of some state employees, some servicemen and women, and some hospital patients with “emotional surveillance technology”. Details of the technology, reportedly developed under the Government funded Neuro Cap programme at Ningbo University, remain secret, but Chinese officials claim it can identify mood, fatigue, concentration levels and, harnessing machine learning technology, predict behaviours such as violent outbursts from hospital patients (Chan, 2018).

Both the technology and the system are defended by appealing to Bentham's arguments. Chinese officials claim that they can boost productivity – specifically, for example, in the case of the Hangzhou State Grid Electric Power Company where profits have reportedly increased by US$315 million since the technology was introduced in 2014. It is claimed it allows a better understanding of workers, to allow them breaks and time off when needed, protects doctors from violent patients, and regulates behaviours (Chen, 2018). Overall, 'sincerity construction' is said to be about building trust. Essentially, the idea is that someone is "once untrustworthy, always restricted", as one of the notices endorsed by eight governmental ministries and the Supreme People's Court announced. In consequence, minor infractions from jay walking to major infractions such as breaking the law reduce one's score and access to services (Reuters, 2018). But untrustworthiness is as much about political opinion as criminality. Social credit regulations are already being used to force businesses to change their language to accommodate the political demands of the Chinese Communist Party, while the surveillance measures ensure everyone upholds discourse that promotes the party-state leadership's version of the truth (S. Hoffman, 2018).

We can see that what is not said by the Chinese authorities is just as important as what is. It seems probable they are telling the truth in arguing that part of their aim is social harmony and productivity. But the system is also an extension of totalitarian control of the country in the interests of the survival and extension of the power of the Chinese Communist party (S. Hoffman, 2017). Surveillance in China threatens to approach or even surpass that described by Orwell in *1984*. The only way to avoid detection of dissident opinions may soon be to suppress dissident thoughts entirely. If anything was going to induce Foucault's reflexive control it is this.

China may have the most advanced state surveillance apparatus in the world but Russia's is Orwellian too. In just one example, Moscow's Mayoral office claims that every one of the city's apartment blocks has the entrance covered by a CCTV surveillance camera, with the state's surveillance equipment supplemented by any number of privately-owned cameras which can be connected, for free, to its cloud based facial recognition system in return for the footage being provided, in real time, to the Government (mos.ru, 2017).

In the more democratic nation of Singapore (ranked 69th 'most democratic' out of 167 nations in the 2017 Economist magazine's democracy index (Economist Intelligence Unit, 2018)) the surveillance network is credited with resulting in a crime rate so low that in 2016 the city-state had 135 days without any crimes at all (Saiidi, 2018). There is so little crime that many new stores are built with no locks on windows and doors. Starbucks in Raffles Place, a major underground station, simply hangs a rope across the entrance when closed. To achieve this, Singapore's authorities installed 12,000 surveillance cameras in 2012 (Saiidi, 2018), added 11,000 in 2016 – ensuring every one of the city's housing board blocks was covered (Sim, 2016) - and set-up a crowd-sourcing scheme for the public to send in video footage to the authorities (Sim, 2016), adding an estimated 1.3 million privately owned fixed surveillance cameras to the network (Webber, 2017). All this, plus all the cameras on smartphones carried by its 5 million+ citizens. As if this wasn't ubiquitous enough, Singapore will soon be adding cameras with facial recognition technology to all 110,000 of its lampposts in 2018 (Aravindan & Geddie, 2018). Again, the arguments of Bentham are to the fore in justifying the expanded surveillance – just feeling watched will be enough to prevent aberrant behaviour (T. Lee, 2005).

## 1. Introduction.

Western democratic states, too, have also extended their surveillance capabilities significantly. In 2011 Big Brother Watch, a civil liberties charity, reported that the average individual in London was recorded by 300 different cameras in a single day (Big Brother Watch, 2012). In 2016 their data showed that London's authorities increased the number of cameras they operate by over 70% (Big Brother Watch, 2016), suggesting that the average Londoner is now filmed by more than 500 cameras a day. Proliferation of CCTV continues apace in the UK, US (Draper, 2018), Europe (Hempel & Töpfer, 2004), and globally (Norris, McCahill, & Wood, 2002). In many urban areas the State's gaze, in public spaces at least, is approaching the panoptic.

We might imagine that CCTV research could help us to understand whether the presence of a camera, and perhaps, therefore, the existence of other 'technologies of control' changes behaviour. However, current research into the effect of CCTV is ambiguous and often contradictory.

A thorough review of all the studies on the effectiveness of CCTV available at the time (2008, updated 2009) revealed that there was, in general, a positive correlation between the use of CCTV surveillance and a reduction in crime with a meta-analytic effect size suggesting the presence of CCTV cameras reduced crime by 16% (Welsh & Farrington, 2008, 2009). Welsh & Farrington found that the use of CCTV in car parks had an even more dramatic effect with crime falling by 51% compared with the control car parks without CCTV surveillance.

A 2017 meta-analysis of what the authors judged to be the most rigorous and best designed studies of surveillance camera's effects on crime published after Welsh & Farrington's (2009) review also found that the presence of CCTV cameras was associated with reductions in crime. The presence of cameras correlated with reductions of 24-28% in public streets and underground railway stations in city centres, and much smaller reductions in unruly behaviour at football matches and theft in retail outlets (Alexandrie, 2017).

Although CCTV effects have been found in meta-analytic aggregations, results within those analyses are variable and seem to be highly context dependent. Both Welsh & Farrington (2009) and Alexandrie (2017) review studies that found no effect, and even some where the installation of CCTV correlated with an increase in crime. Alexandrie's results for instance, partially contradict and partially corroborate Farrington & Welsh's finding on the increased effectiveness of CCTV in enclosed areas. Although Alexandrie found no CCTV effect in car parks, or in underground train stations in the city centres, he does find that installing CCTV correlates with crime reduction in suburban underground stations further from the centre of cities (Alexandrie, 2017). Perhaps Alexandrie's findings add nuance to Farrington & Welsh's, indicating possible boundary conditions for CCTV's effect. Or perhaps they indicate that CCTV's effects have been over-estimated or simply do not exist, with previous findings being false positives caused by other environmental factors.

For example, evaluation of insurance and police data in Surrey, a city in British Columbia, Canada, found a significant drop in car crime in car parks over a 12 month before and after study, but not when considering the data over a longer period of three years and with a wider control area (A. A. Reid & Andresen, 2014). It might be that crime reductions found in other cities where CCTV has been installed are down to such annual fluctuations in crime, rather than being caused by CCTV. Similarly, the results of researchers exploring the effect of a huge increase in CCTV camera use in Lyon, Saint-Etienne and Grenoble, in France, found the presence of cameras had little or no effect on crime figures (Germain, Douillet, & Dumoulin, 2011). This might be down to cultural variation in CCTV's effects, to other regional or environmental effects or, again, indicative that other studies have found false positives.

## 1. Introduction.

It is not that there have been no attempts at finding other outcome measures beyond reported crime to identify whether CCTV exerts a deterrent effect – but rather it is that these are also vulnerable to confounding factors. The insurance data used in the Canadian study mentioned above is one such example (A. A. Reid & Andresen, 2014). Two further examples illustrate the point well. The first measured the number of people reporting to accident and emergency rooms with assault and violence related injuries in five English cities that installed CCTV when compared with cities that did not do so (Sivarajasingam, Shepherd, & Matthews, 2003). The second reviewed incidents of anti-social behaviour from referees' reports on the number of objects thrown on to the pitch (such as coins, bottles and lighters) at Swedish football stadiums (Priks, 2014). Both found a positive effect from the presence of CCTV based on a reduction in, respectively, reported injuries and reduced anti-social behaviour. But even with alternative outcome measures, field experiments are always vulnerable to the charge that something else may have caused any effects found. Given the proliferation of CCTV cameras and other surveillance technologies, the need for more rigorous laboratory research is clear.

Militaries, too, are exploring the possibilities made available to them from access to modern surveillance capabilities which are producing unprecedented amounts of data. Retired US Navy Commander Ted Johnson and retired United States Air Force (USAF) General Charles Ward explained in an article last year that in 2011 the USAF produced full motion video footage with a total running time of some 37 years, taken from satellite collection, drones and other aircraft – 2011 being the only year for which any estimate of the scale of USAF collect has been released (T. Johnson & Wald, 2017). The same authors, describing the constellation of surveillance capabilities available to the US military, argued that the US should aim for 'operational omnipresence', suggesting that:

## 1. Introduction.

‘*Today, cyber- and satellite-enabled surveillance, coupled with traditional forms of intelligence gathering and the ubiquity of the press and personal devices, means that* ***a global electronic version of the panopticon is possible****… …What was thought to be in secrecy is now under a spotlight. From that point forward, …[any]… nation would assume it is being watched. This, too, is a form of presence, and the essence of perceived presence’* (Wald & Johnson, 2017) [emphasis added]

Beyond CCTV and drone video footage, Edward Snowden’s revelations have shown the extent of state surveillance capabilities in the US and UK. I will not comment further on this given my professional responsibilities but for the interested reader Snowden’s disclosures can be found in the report *NSA Files: Decoded*[1] as presented from the perspective of the Guardian newspaper, the first publication to break the Snowden story, in a comprehensive summary by journalists Macaskill and Dance (2013).

Some research has suggested that these revelations may have caused Foucauldian self-censorship, or ‘chilling effects’, as this is now often termed. Penney (2016) found that traffic to Wikipedia articles on subjects associated with privacy concerns saw an immediate and statistically significant decrease of 30% in the month after the Snowden revelations broke in 2013 with the trend continuing for the remainder of the month. Further research has shown that Google users searches for terms that might get one into trouble with the US government, or embarrass one if a friend were to know what one had searched for, dropped significantly in the month after the Snowden revelations (the terms selected were determined independently by the ratings of participants in a preliminary study on the online research platform MTurk) (Marthews & Tucker, 2015). This reduction was greatest in countries that were diplomatically closest to the US, and which had lower levels of state surveillance

---

[1] https://www.theguardian.com/world/interactive/2013/nov/01/snowden-nsa-files-surveillance-revelations-decoded#section/1

(where state surveillance was defined by the proxy measure of whether citizens are mandated to carry a state ID card).

The authors admit that their data does not allow firm conclusions to be drawn from this finding, but suggest the difference may be because citizens in those states with high surveillance levels or which are obvious targets of US surveillance already self-censor, and consequently the Snowden revelations had an effect only on those who previously felt anonymous on the internet as they began to self-censor now that they too knew they might be watched (Marthews & Tucker, 2015).

Relatedly, self-reports from journalists suggest that as many as 28% curtailed social media activity post-Snowden, 24% claimed to have avoided certain topics in phone or email conversations, and 16% claimed to have avoided speaking about a particular topic altogether (FDR Group & PEN American Center, 2013).

Taken together, then, Marthew's & Tucker's analysis of Google searches, their exploratory finding of the greater effects in countries where citizens usually feel anonymous online, Penney's (2016) Wikipedia findings, and the aforementioned self-reports of journalists might all be seen as providing some evidence for Foucault's notion of reflexive control in the 21st Century context. Thus, when considered alongside the technological developments in China, Russia, Singapore, and generally across the globe, we can see that in authoritarian states and democracies alike we have cause to feel more and more watched…and cause to understand precisely what effect this might have on our behaviour.

### 1.1.3. Surveillance Capitalism

As the aforementioned discussion demonstrates State surveillance is becoming increasingly ubiquitous, and as such, the urgency of understanding the concerns of the philosophers may never have been greater.

# 1. Introduction.

And this is before we take account of the rise of 'surveillance capitalism' a business model predicated on watching everything we do and manipulating and selling the data collected. This has led to far more extensive surveillance in democratic countries and worldwide being undertaken by private companies to such a degree that in 2010 Facebook CEO Mark Zuckerberg was reported to have said that 'the age of privacy is over' (Kirkpatrick, 2010).

His comments had been slightly twisted. In fact, Zuckerberg claimed only that Facebook had changed social norms around privacy (TechCrunch, 2010). Just 3 days later his interviewer, the leading tech commentator Michael Arrington, wrote on his blog that "*The point is, we don't really care about privacy anymore. And Facebook is just giving us exactly what we want*" (Arrington, 2010). Others in the online commentariat noted that this change in social norms was hardly new, observing that "*The age of privacy did not end in 2004 with Facebook and the social web. Equifax, Transunion, Capital One, American Express and their cousins raped our privacy*" (Lindzon, 2010). The truth, of course, is that privacy's death had been forecast for some time. In 1999 Scott McNealy, CEO of Sun Microsystems, told the world bluntly: 'Privacy is dead. Get over it' (Couts, 2012).

This loss of privacy is a consequence in large part of what has become known as 'surveillance capitalism' in which technology companies ask customers to exchange privacy for services (Zuboff, 2015). For example, if you have your smart phone with you as you read this, you've been geo-located approximately 23 times in the past minute, while in the next hour you'll have been located nearly 1400 times (Howard, 2017). Similarly, Facebook knows where the vast majority of its 1.94bn active users - defined as those that have logged on in the past month (Statista, 2017) - live, work, and socialise, and also what time they get up, when they went to bed, where they went to bed, and, by inference, perhaps even who with.

## 1. Introduction.

Then there is Google. Google tracks the things you search for, the websites you visit, the videos you watch, the ads you click on or tap, as well as your location, device information, and, through Gmail, the emails you send and receive, their content, the contacts you add, your calendar events, the photos you upload, the documents and material you save on your Google Drive, and if you have an account: your name, your birthday, your gender, your phone number, and your nationality (Google, 2017).

Google is *listening* to you too. In 2016 it was widely reported that the conversations of anyone using Google voice search on their phone had been recorded - not just the searches, but the actual conversations people had in the vicinity of the device on which the app was running, in order to improve its language recognition and Google's targeted advertising (Griffin, 2016). Similarly, if you talk to Apple's Siri, your instructions are saved for 2 years, and your data, and recordings, may be shared with third parties. Samsung operates a similar policy. The whistle-blower, *Fallenmyst,* posted on Reddit in 2015 that he had been working for an app developer rating the accuracy of text interpretations from recorded naturalistic speech when he noticed a pattern and realized he was listening not to random recordings but to users' real-life interactions with their smart phones. He described how even the most intimate conversations were passed to him, writing: "I heard everything from kiddos asking innocent things like 'Siri, do you like me?' to some guy asking Galaxy to lick his butt. I wish I was kidding" (Cluley, 2015). That was you? Well, now you know. There is little that goes unobserved.

But there are already a number of companies whose surveillance is far more extensive and intimate than that described. Nielsen company eXelate surveils your activity online and can determine your age, sex, ethnicity, marital status and profession, as well as your current consumer interests, matching all this with offline purchasing data through partnerships with supermarkets and other retailers (Nielsen, 2017a, 2017b). Then we have Towerdata.

## 1. Introduction.

Towerdata claim to have 80% of U.S. email addresses, all the same data on you as eXelate, but also to know your income, whether or not you have children, your postal address, education, your name, your telephone number, your homeowner status, your length of residence, the current value of your home, your interests (including what you create, like, view and interact with, as well as your connections on social media and interactions with those contacts)… "and much more" (Towerd@ta, 2017).

Furthermore, the internet of things[2] is set to remove privacy from the heart of your home, blurring the online/offline boundary, enabling analysis of yet more personal data still. In 2014 Google bought Nest (Wohlsen, 2014), a US company that builds devices to create 'a thoughtful home: a home that takes care of the people inside it and the world around it' (Nest, 2017b). This thoughtful home helps you turn your heating on remotely, perhaps as you set off back from the office or, if you prefer, as you get within a certain distance of home. It learns what temperatures you like in different rooms at different times. Its smoke alarms can talk to you – so you can shout when you've burnt the toast to stop them going off. Nest's security, fire alarms and apps also join forces with your lights and your locks – lighting the way to help you get out in a fire, for example, or flooding the room with light to scare off an intruder at the window. In addition, it reminds you not to forget to switch the alarms on as you leave, and saves power by turning the lights on and off as you move around the house. But, of course, all this extra security comes at a price. Google knows when you get in and when you go out. It knows what rooms you are in. It knows not just when you put your phone down and stop using social media, but when you cook, when the lights are turned on, turned off, turned up, turned down…you name it, allowing all sorts of inferences to be made as to what you might be doing.

---

[2] **Internet of things** *n.* a proposed development of the Internet in which many everyday objects are embedded with microchips giving them network connectivity, allowing them to send and receive data. Oxford English Dictionary. Retrieved 22 September, 2018, from: http://www.oed.com/viewdictionaryentry/Entry/248411.

# 1. Introduction.

Think that's the end of it? Think again! In fact, it's just the start. The plan to build a home that is not just thoughtful, but 'conscious', in the words of Nest's founder (Wohlsen, 2014), to do this would mean building an aggregate picture of your behaviour – providing unprecedent insight into you and your activities, removing privacy not just from what you have done, but projecting forward to know what you will probably do in the future. For example, Nest's Cam IQ security cameras use facial recognition software via its 'supersight' lens to know who is calling (Nest, 2017a). They might add this to data on your water and electricity usage, for example to know how you act before guests arrive, when you boil the kettle, wash your clothes, take a shower, or flush the loo. Google, or analytic middle-men like eXelate and Towerdata, might then link all this to when you order your online shopping, what you order and depending on data sharing agreements around your loyalty card, maybe when and what you buy in store too. With all this information, Google (or their rivals – they are not alone in pursuing these technologies) will know who you're meeting, when, how different guests at your home are fed and watered and might infer what you're doing at different times and what you need done next as a result. And Nest isn't the be all and end all of this. Far from it. Professor Philp Howard, Director of the Oxford Internet Institute predicts that by 2020, when there will be eight billion people on the planet, there will be three to four times as many internet-connected devices (Howard, 2015). The Organisation for Economic Cooperation and Development in Europe predicts that the average family of four will have fifty internet connected devices by 2020, up from an average of ten today, and engineers seem to think there could be far more – anticipating the arrival of the internet of things by expanding the web to allow for an additional $2^{128}$ new unique URLs (Uniform Resource Locators): enough for every atom on the planet to have 100 *www.* addresses each (Howard, 2015, pp. xii-xiii).

Once such development that is with us already is the importing of virtual assistants to our living rooms, and other locations within our homes (Brenner, 2017; Goldstein, 2017;

Henderson, 2017; Waddel, 2016). These devices are either standalone pods or applications within your phone, and they have various privacy settings, but all record to some degree. In November 2015, when Georgia Police Office Victor Collins was found dead in a hot tub in the garden of his friend, James Bates' home in Arkansas, the police subpoenaed for the records of the Amazon Echo that had live streamed the music that Collins had been murdered to. Noting water utility records showed 140 gallons of water had been used at Bates' home in the middle of the night after the murder (more surveillance) and that Bates had made calls during the night – when he claimed to be asleep - they guessed that Amazon had recordings perhaps of what happened before, during and after the murder. Amazon stalled at first, presumably fearing the damage that would be done when people realized just how much intimate data they, and the device, actually held. But in February 2016, when Bates indicated he was content to hand the device over, they complied (Gallagher, 2016; Grant, 2017).

The case, it transpired, was eventually dropped by the County Prosecutor (Chavez, 2017), so we will never know what the Echo and Alexa's witness testimony might have revealed. However, the very fact that Alexa can be a witness is revealing enough for us to begin to comprehend just how ubiquitous surveillance has become and how much more extensive it is set to be.

And if just feeling watched can change behaviour, as Bentham argued, and if it can enable 'reflexive control', then there is an urgent need to understand the psychological effects of this so-called 'death of privacy'.

### 1.1.4. Surveillance, Sousveillance, Activism & Policing

Just as burgeoning surveillance capabilities are empowering the state, so protest movements are adopting 'sousveillance' (under sight) as a counter to surveillance (oversight) (Mann, Nolan, & Wellman, 2004). In this conception, pointing cameras back at the authorities offers the citizen a chance to exert reflexive control in reverse, forcing police and state authorities to act within the law and accepted norms for fear of exposure (Mann & Ferenbok, 2013; Sandhu, 2016). Sousveillance has been used to try to moderate behaviour in many situations, but three examples suffice to make the point – in the prosecution of police officers for abuse of powers during the G20 summit in Toronto, a move which has forced police services worldwide to consider the possibility they might be filmed (Mann & Ferenbok, 2013);  in the use of dashboard cameras to counter corrupt police seeking bribes in Russia (Balmforth, 2012); and in the Black Lives Matter campaign to expose racist and violent behaviour by US Police (Zuckerman, 2016).  In all three examples this 'under sight' intervention has indeed been successful, to some degree, in helping to constrain the actions of the state.

Sitting somewhere around the meeting point of the surveillance and sousveillance debate rests the issue of body-worn cameras on police.  Randomized controlled field trials have yielded positive findings for the effectiveness of body-worn cameras on policing.  Those wearing them on a trial in California, for example, were half as likely to use force in comparison with those who were not wearing them, and complaints against those wearing them were ten times lower than complaints against those who were not (Ariel, Farrar, & Sutherland, 2015).  The authors suggest the cameras achieve these dramatic results by heightening self-awareness.  To corroborate this finding, a follow-up study of police in the US and UK found police use of force fell 37% when cameras were always on, but rose 71% when officers could turn the cameras on or off at their discretion (Ariel et al., 2016).  In a further follow-up study, the largest randomised trial of police body-worn cameras to date, complaints against the police fell by 93% for those wearing the cameras in comparison with those who were not (Ariel et al., 2017).

In fact, research by the UK's Behavioural Insights Team suggests that the effects of body-worn cameras go further than simply attenuating hostile behaviour. Those police officers wearing the cameras in their 2016 trial took an average of 3.3 fewer days off, 20% less than those who did not wear them. More significantly, those wearing the cameras seemed to be more discriminate in selecting those to submit to stop & search procedures, with 29.9% of stop and search encounters resulting in a positive outcome (e.g. a possession charge), a statistically significant increase compared with 25.6% of encounters amongst those not wearing the camera (The Behavioural Insights Team, 2017).

In this body-worn camera research, the presence of the cameras would seem to regulate the behaviours of both the police and the public with whom they were interacting, with both groups conforming more closely to social norms. Such being the case, the counter-insurgency expert, former advisor to General Petraeus in Iraq and security consultant David Kilcullen describes a related and intriguing idea. Kilcullen suggests that the increasing ubiquity of surveillance technology enabled the Arab Spring, the series of uprisings from Tunisia, via Egypt, to Libya and Syria, that broke out in 2011:

> '*I think this is about more than just the convenience of electronic connectivity though. Constant access to the digital world, letting people upload images or tweet what's happening to them, creates a sense of security. There's always an actual or potential witness to what's going on: someone's watching, ready to blow the whistle if the authorities pull something brutal or repressive. It's as if there were always a media crew of reporters and cameramen watching out for you – but a virtual, digital, distributed crew enabled by constant connectivity. This idea of "Web as witness" – the protection that comes from virtual monitoring by independent outsiders, and the restraint this imposes on governments – is the flipside of the privacy concerns that go with our ever-connected environment. In a sense, it allows remote actors to extend their normative systems into places where they can't*

> *physically be. This idea of the permanent, universal witness is a new element in conflict, politics, and human rights advocacy alike, it's entirely an artefact of the connected, urban world and it's mostly a good thing*'. (Kilcullen, 2013, p. 246)

Indeed, activists themselves have discussed in online forums the import of YouTube and streaming sites as witnesses to their protests. In Syria there were tens of thousands of webcams and mobile phones used to record and broadcast protests against the Assad regime in 2011-12 (Nachawati Rego, 2012; Shanley, 2012) with many efforts centrally coordinated by the Syrian Observatory for Human Rights (Wikipedia, 2018) – whose eye-like logo (*Figure 1*) is suggestive of its sousveillant purpose - as well as by online media curation organisations such as Ugrait News, Sham Network News and Homs Network News (Starr, 2012, pp. 61-62). Indeed activists have been able to enlarge Syria's internet in geographic area and number of users during the country's 7 year civil war despite the devastation of the country's infrastructure, and have done so in part because they have been desperate to be watched (Waters, 2018).

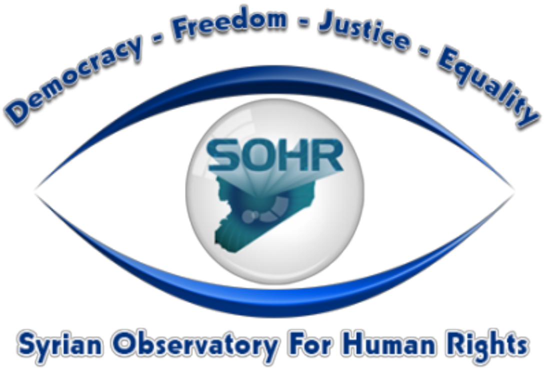

Democracy - Freedom - Justice - Equality
SOHR
Syrian Observatory For Human Rights

*Figure 1.1. Logo of the Syrian Observatory for Human Rights*

Of course, we must acknowledge that Kilcullen's explanation does not rely solely on the idea that it was the feeling of being watched that empowered or constrained behaviour in this situation. It might be that the effectiveness of the web as witness came more from the threat of international sanction against the state and its agents, and the belief on the part of the protestors that their visibility made them more likely to receive the support of outside powers. It must also be noted that more sceptical takes of the effect of internet activism suggest that it isolates activists who can then be more readily identified by repressive states, and may, at the same time, enable superficial signalling - ‘clicktivism’ - at the expense of meaningful committed organisation. Similarly, critics suggest that web-based surveillance is more cost than benefit to activists in other ways. That it may simply entrench existing power structures through commercially mediated online fora such as Facebook - enabling corporate entities to dispossess us of our data and use it to sell goods back to us with ever more sophisticated techniques, to control the media we see, to shape our preferences and use this power to watch, regulate and perhaps discourage or redirect our interests, to influence the groups we see via search engines and social media, to encourage us to sympathise with some and not others, and thus which groups we support and join (Fenton & Barassi, 2011).

Whatever the truth of the matter may be, what is *not* in doubt is that many of the arguments being advanced suggest an urgent need for more empirical evidence around the effect of surveillance on behaviour. Similarly, although we must acknowledge that in some of these examples, such as the case of police body-worn cameras there clearly are potential consequences: in this case for both the police and those they interact with should either transgress legal norms. However, there is again a case to be made in this instance, and in the other examples here to suggest that Bentham’s panopticon may be operating through modern technology. The problem, however, is that it is not possible to separate out subconscious or non-coercive surveillance effects - such as the role of increased self-awareness induced through feeling watched from the effect of explicit concerns regarding future prosecution.

Our studies aim to address this conundrum directly, seeking, amongst other things, to identify whether heightened self-awareness is in part responsible for cueing police and public to behave differently when they feel watched or whether it is only explicit concerns regarding the future consequences arising from being watched that make the difference.

### 1.1.5. Surveillance, Peace-Keeping & Humanitarianism

The most immediate motivation to conduct this study comes from my experience on peacekeeping operations. In 2007-08, serving as a UN Peacekeeper and the information officer in the Gali Sector within the disputed territory of Abkhazia with the United Nations Mission in Georgia, I proposed that the UN purchase its own air surveillance radar, coastal surveillance radar and unmanned aerial vehicles. The idea was to make the feeds from this surveillance equipment available live on the internet. UN peacekeepers are often described as the eyes and ears of the world. On what the UN Charter calls *Chapter 6 Peacekeeping Operations* (United Nations, 1945) peacekeepers are unarmed, able only to report what they see and hear. Consequently, my proposal would have effectively crowd-sourced thousands more peacekeepers – yet more eyes and ears - from around the world at a stroke. Within the mission, we speculated that it might have an additional effect. Perhaps having the world watching so directly would itself reduce cross-border incidents and breaches of the ceasefire agreement.

This idea has since been tested. On 29 December 2010 the actor George Clooney and human rights activist John Prendergast launched the Satellite Sentinel Project (SSP). The SSP crowd-sourced imagery analysis much as I had proposed doing with the UN. From 2011 to 2015 the SSP constantly reviewed and regularly published on the internet satellite imagery of Sudan and South Sudan. The imagery was online for all to see with the tagline 'the world is watching because you are watching' and the explicit aim of deterring war and human rights abuses in the region. In a press release Clooney said of the project: "We want to let potential perpetrators of genocide and other war crimes know that we're watching, the world is watching… …War criminals thrive in the dark. It's a lot harder to commit mass atrocities in the glare of the media spotlight." (Satellite Sentinel Project, 2010). Here again we run into the ideas of Bentham, using surveillance to deter bad behaviour.

A critical analysis of the project noted several limitations that had reduced its effectiveness in preventing violence, but still cited a number of successes: identifying mass graves while looting and killing was ongoing, identifying forces moving to attack civilian areas and food convoys, and, above all, forcing the world to bear witness to what was happening even if not provoking direct international action (Raymond, Davies, Card, Achkar, & Baker, 2013). However, notable by its absence is any investigation of Clooney's claim that just bringing the light of surveillance, the eyes of the world to the conflict, might reduce war crimes.

A similar programme launched in June 2007 by Amnesty International USA, "*Eyes on Darfur*", aimed to use crowd-sourced online activists to 'watch' 12 villages in Darfur via imagery satellites, again with the aim of deterring and reducing human rights abuses. Here, however, a detailed quantitative evaluation of the six-month project taken against a baseline of violence across the region found that in fact violence in these villages had increased by an additional 15% during the intervention and a further 20% after it ended (Gordon, 2016). In a sense this is not surprising – where the strategic logic favours more violence, the 'subconscious cueing' would likely be limited in its effects. In field interventions such as these the number of confounds is such that we cannot rest too much weight on the conclusion of just two examples. But as large-scale surveillance programmes predicated on the ideas of Bentham they clearly demonstrate the need for more controlled research into the effect of surveillance on behaviour before further such interventions are attempted.

Although the limited effectiveness of the Satellite Sentinel Project and Eyes on Darfur interventions suggest otherwise, the argument that such perceived surveillance and monitoring are able to change global norms and reduce abuses remains prevalent (Risse et al., 1999). David Whetman of King's College, London argued in the International Journal of Human Rights in 2016 that 'flying cameras', that is, unmanned aerial vehicles with full motion video or other optical sensors, could be rapidly deployed by the UN anywhere on the planet and that if belligerents were aware of the deployment 'the fear of being observed may be enough to modify behaviour' (Whetham, 2015).

Added to this, the United Nation's burgeoning aerial surveillance capabilities have been justified on the grounds that '...if criminal/violent elements are aware that the United Nations can operate silently and without detection, a powerful deterrent is created, instilling fear in violators even when aircraft are not present' (Dorn, 2014, p. 121). This idea, that just watching can help bring peace, is present within the justification for the aforementioned United Nations' *Chapter 6* missions, where unarmed peacekeepers are deployed as 'the eyes and ears' of the international community with no mandate to intervene (Sisk, 2017, p. 493).

And surprisingly, perhaps, this kind of peacekeeping works. Indeed, the most thoroughgoing quantitative examination of UN peacekeeping missions of which I am aware suggests it can be dramatically effective. Columbia University Professor of Political Science Virginia Fortna ran a series of statistical analyses looking at the hazard risk, a ratio measure similar to the odds ratio, while controlling for diverse variables such as the number of deaths in a conflict, whether there had been a decisive victory for one side, infant mortality, terrain etc. Her various models consistently include around 100 different peacekeeping missions. Even in her most conservative estimate Fortna finds, *ceteris paribus*, that the presence of UN peacekeepers halves the chances of renewed war. Her less conservative estimate suggests that it reduces the chances of war by as much as 75-85% (Fortna, 2008, p. 125). One cited causal mechanism for just *how* such peacekeeping can be so effective is the idea that simply watching is enough to change behaviours (Diehl, 1993, p. 10; Doyle & Sambanis, 2000, p. 785; Rotberg, 2000, p. 5).

Fortna herself gives the idea that surveillance alone is enough short shrift (Fortna, 2008, pp. 76-103), but the explanation for how peacekeeping might work offered by other scholars reads a lot like Bentham's panopticon applied to peacekeeping. And although the limited evidence from the Satellite Sentinel Project and Eyes on Darfur studies suggest the power of surveillance alone is not enough to offer panoptic benefits, nor to allow Foucauldian reflexive control, the debate continues, with some humanitarians arguing for an 'Ambient Protective Effect (APE)' provided by increasingly ubiquitous surveillance of disaster areas by information communication technologies, and others arguing the contrary (Sandvik & Raymond, 2017). Since neither side has direct evidence to support their claims, the research presented in this thesis aims to either corroborate or refute these contentions empirically.

Moreover, in the preceding discussion, in order to understand the implications of surveillance in different contexts I separate out state, commercial, policing, military and activist surveillance into separate domains. However, it is important to note that some surveillance scholars would likely resist this conceptual separation. The panoptic effects of modern surveillance are, it is suggested, created by 'liquid surveillance' in which we often cannot know whether we are being watched by the state or other entities, and do not know whether the data collected by one might be put at the disposal of another (Lyon & Bauman, 2013; Manokha, 2018; Zuboff, 2015). If we accept this premise, the urgency of understanding whether there are effects on behaviour solely from feeling watched becomes more urgent still.

### 1.1.6. So What?

In approaching the problem, it is important to get beyond pitching Bentham and Foucault's arguments against each other. As can be seen from the evidence presented, qualitative methods and quantitative correlational studies of real-world data have not been able to provide sufficient insight. How then might we approach providing the evidence needed to better understand surveillance effects?

## 1. Introduction.

Psychological science offers experimental techniques that allow researchers to manipulate surveillance conditions, to limit potential confounds, and provide objective analysis of outcome measures. Such experimental psychology also allows us to carefully describe our designs, such that they can be objectively critiqued, our results replicated (or not), and our findings challenged on an equally empirical and objective basis.

With the approach decided upon, two possible hypotheses arising from the philosophical discussion and contemporary context were advanced for exploration:

H0. Surveillance has no effect on behaviour unless coupled with an explicit consequence.

H1. Perceived surveillance – just feeling watched – can change behaviour.

These two hypotheses, together with the ubiquity of surveillance in general, generate a number of societal implications of both a practical and an ethical nature. Under the null hypothesis (H0), for example, we might be less concerned by the growing ubiquity of surveillance capabilities *per se,* providing that regulation prevents the specific data from being exploited to ill-effect - intentionally or otherwise. Furthermore, if surveillance alone causes no psychological consequences - the adage 'only those with something to hide have something to fear' holds. The rest of us will be unaffected and need not worry.

In contrast, if H1 holds, we need to think of the practical implications of any such effect, both negative and positive. In terms of the negative, we need to think much more carefully about the consequences for society of the ongoing, rapid, and seemingly exponential increase in surveillance capabilities across the commercial, governmental, and social domains. In terms of the positive, we need to consider where those capabilities might be applied to best effect, such as in peacekeeping, activism, and law enforcement, with a view to constraining both criminal behaviour and government excess.

## 1. Introduction.

Surveillance, then, could be a double-edged sword –you can't enjoy its benefits without incurring its costs and vice versa. If it increases conformity and causes self-censorship, it should be regulated and controlled. If it has the potential to increase productivity, trust, safety and social cohesion – maybe even to help bring peace to troubled areas of the world – then we ought to expand it. The arguments of Bentham and Foucault are still, to this day, being pitched one against another as they were throughout the last century, while the fears that Orwell illustrated so brilliantly are rarely far from the public discourse. Balancing the promise and perils of surveillances will need empirical evidence to be weighed carefully if it is to be done effectively – pitching Bentham's utopian surveillance dream against Foucault and Orwell's dystopian surveillance nightmare has taken the debate as far as it can, helping us to define the competing hypotheses.

However, differences aside, Bentham and Foucault's arguments both have something in common. They rest partially on the premise that surveillance has panoptic effects beyond the immediate and explicit concerns as to the *consequences* of being watched. What is urgently needed, then, is an evidence-based assessment of the central argument in both claims. Without it, the discussion will continue to be based on contrasting, competing, but ultimately unempirical assertions. On untested optimism against speculative, unexpurgated fear.

1. Introduction.

## 1.2.The Psychology of Surveillance: A Literature Review

### 1.2.1. Surveillance Studies & Psychology

The proliferation of surveillance described in Chapter 1 has helped spark the creation of a new academic field of Surveillance Studies which has been described ‘as a cross-disciplinary initiative to understand the rapidly increasing ways in which personal details are collected, stored, transmitted, checked, and used as a means of influencing and managing people and populations’ (Lyon, 2002, 2007; Lyon, Ball, & Haggerty, 2012). How surveillance changes human behaviour and what effect it may have on individuals, groups and societies are pressing questions. Yet, in spite of claims that Surveillance Studies is an interdisciplinary approach there remains very limited psychological research on the subject (Ellis, Harper, & Tucker, 2016).

This lacuna in psychological research is important not only in relation to public policy but also for the advancement of psychology itself. The debate around the first failed replication of Zimbardo’s famous (although now perhaps discredited (Blum, 2018)) Stanford Prison Experiment highlights the import of surveillance research in psychology. Zimbardo suggested that ‘good people’ divided arbitrarily into guards and prisoners can quickly fall into patterns of abuse and submission (Zimbardo, 2011). However, a BBC commissioned 2002 study by British psychologists Haslam and Reicher had different results. In Haslam & Reicher’s replication the prisoners identified with their role but the guards did not. This made the guards reluctant to impose their authority and the prisoners eventually overcame them. Participants then established an egalitarian social system (Reicher & Haslam, 2006). This seemed to fundamentally undermine Zimbardo’s findings.

In an extensive rejoinder, Zimbardo argued that this was in part because:

> "*In our study, video-recording was concealed and never apparent to participants, and none wore microphones to make them aware of being under surveillance. In theirs, it was obvious that everything was being recorded at all times.*" (Zimbardo, 2006)

Responding in turn, Haslam & Reicher (2006) downplayed the significance of the presence of cameras on participants but conceded that:

> *"…even where the cameras did have an impact, this does not invalidate the study. Instead, it simply means that it is necessary to clarify the processes underlying their impact in order to understand the wider implications of particular findings. Along these lines, Zimbardo acknowledges in his postscript that it is far from trivial to conclude that surveillance may have affected the willingness of the guards to impose their power.*"

Amongst Haslam and Reicher's recommendations in this retort to Zimbardo's criticism was that the effect of surveillance on behaviour needed further examination. Sixteen years later, in the aforementioned review of psychology's contribution to surveillance studies, Ellis et al. (2016) were writing in the same publication arguing for its import and calling for exactly the same thing again.

The furore around a recent (2018) article published on the British Psychological Society's *Psychologist* blog shows just how important, and little understood, the psychology of surveillance remains. The science journalist Jesse Singal (2018) reported on another high-profile psychological study, this one by Strack, Martin, and Stepper (1988), that had found that smiling or frowning provided 'facial feedback' that lifted or depressed mood. Strack et al.'s study failed to replicate in a large scale pre-registered replication attempt in 2016 (Wagenmakers et al., 2016), only to apparently replicate when re-run in 2018 (Tom Noah, Yaacov Schul, & Ruth Mayo, 2018). The main criticisms of the 2018 replication were statistical (Singal, 2018; Tullet, 2018), but central for our purposes is the part of the defence that Strack launched of his study after the original failed replication (Strack, 2016). This was the idea that Wagenmakers et al. (2016) had not found an effect because, as Strack put it, they had differed from the '*original study by directing a camera on the participants*' explaining that '*based on results from research on objective self-awareness, a camera induces a subjective self-focus that may interfere with internal experiences and suppress emotional responses'* (Strack, 2016). Noah et al. (2018) sought to investigate whether Strack's defence was sound. Beyond the specifics (and criticisms) of Strack's experiment, the discussion shows the continued import and relevance of psychological research into the effect of surveillance on behaviour.

### 1.2.2. Watching Eyes Effect

One area of psychological science that has sought to understand the effect of surveillance on behaviour is the investigation of the so-called 'watching eyes' effect. This research investigates whether the mere presence of pictures of eyes or stylized eye images (hereafter: eye cues) is enough to cause us to adjust our behaviour, making us more generous and prosocial (Bateson, Nettle, & Roberts, 2006; Burnham & Hare, 2007; Haley & Fessler, 2005) or less likely to break rules or to behave antisocially (Bateson et al., 2015; Nettle, Nott, & Bateson, 2012). Eye cue research offers a way to study the effect of perceived surveillance on behaviour independent of deterrence effects or behaviour-modifying reputational concerns as might occur when someone *is* being watched rather than just conceiving themselves to be so. The 'watching eyes effect' suggests that just feeling watched may be enough to make us modify our actions independent of deliberative, explicit, conscious, evaluation of the costs and benefits of an action. Perhaps the talismans and charms of the ancients offered more than just superstitious protection (see Figure 1.2. overleaf).

Why might eye cues work? Before we explore the specific experimental evidence in relation to the watching eyes effect it is helpful to consider a theoretical framework within which we can explain and understand why eye cues might have an effect on behaviour. If eye cues do have an effect on our decisions and behaviour, it seems reasonable to assume they are acting as false cues to actual gaze, triggering the normative response to being watched. A reasonable assumption, therefore, would be that we can utilize evidence and theoretical developments from work that examines real, human-to-human, eye-to-eye gaze, and perhaps draw inferences from the responsiveness to gaze and gaze cues in other animals. We start this exploration with the available evolutionary & ethological evidence, then look at the cue itself, the structural design of human eyes in comparison with those of other animals, before turning to human-specific psychological and neuroscientific evidence, to evolutionary psychology, and finally to surveillance specific psychological research.

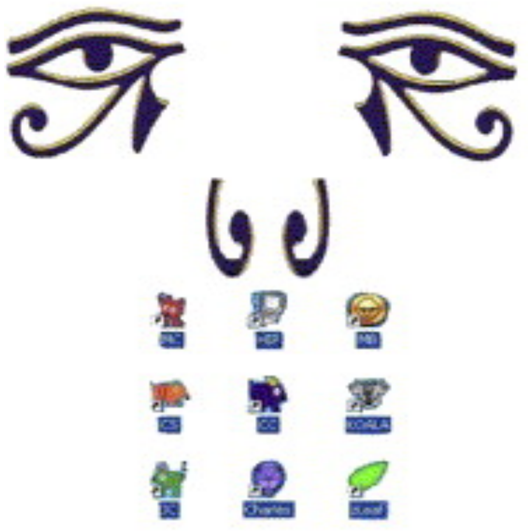


*Figure 1.2. Stylised Eye of the Egyptian God Horus on the screen of Haley & Fessler's (2005) economic (dictator) game – where the presence of eyes correlated with increased participant generosity.*

### 1.2.3. Gaze in Evolutionary Perspective

Evolutionary psychologists understand the mind to be a confederation of 'modules' shaped by evolution to solve adaptive problems endemic to our hunter-gatherer ancestors (Tooby & Cosmides, 1997, pp. xv-xvi). From this premise, theories of adaptive function inform our understanding of how these modules came to be and what affect they might have on human behaviour in the modern environment. Such mechanisms should solve narrowly identifiable sets of biological information-processing problems defined by Darwinian natural selection operating within the context of a species' ancestral mode of life. For humans this means primarily the world of ancestral hunter-gatherers, including pre-linguistic human societies, and members of the genus *hHomo* that predated us as *sapiens*. Further back still it includes pre-hominid primates (Tooby & Cosmides, 1997, pp. xv-xvi) and even the mammals and other animals in the upper reaches of the *Homo* phylogentic tree. Perhaps then we can gain insight into whether eye cues might influence human behaviour from a review of the way other animals respond to gaze and gaze cues – looking for mechanisms that might have existed in our evolutionary ancestors, or selection pressures that might have made an eye cue response adaptive in other animals and perhaps, therefore, in humans too.

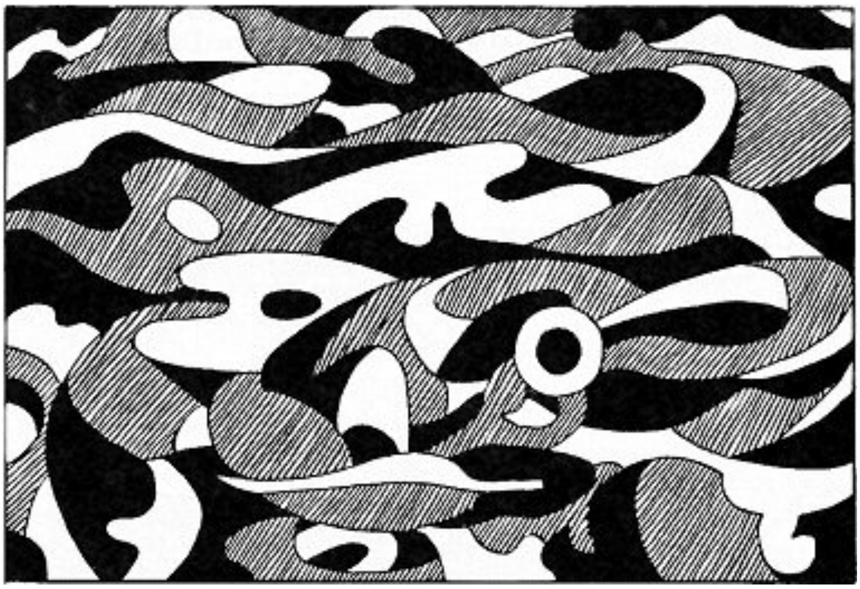
### 1.2.4. Evolutionary Evidence Part 1 Ethology

Looking at the ethological literature what we find are a great many examples of hyper-sensisitivity to gaze and to eye cues in many animal species. For example, ethologists identify *tonic immobility* – the tendency to play dead - as a widespread threat response in many animals when gazed at – and find evidence to suggest this effect can be triggered by eye cues. This has been observed, for example, in crustaceans (crabs) (Brien & Dunlap, 1975), chickens (G. G. Gallup, Nash, & Ellison, 1971), lizards (C. Hennig, 1979; C. W. Hennig, 1977) and hog nosed snakes (Burghardt & Greene, 1988). Further evidence comes from animals that display avoidance responses when stared at. Chickens both freeze when stared at and show avoidance (a reluctance to approach and wariness) of static models of predators but only when the predator's eyes are unobscured – cover up or remove the 'eyes' and the wariness disappears (Scaife, 1976a, 1976b). Chickens are also wary of paired eye cues but not individual ones – perhaps reflective of the fact that their primary predators are mammals such as foxes with forward facing paired eyes (Scaife, 1976a, 1976b). Similar avoidance behaviour has been observed in jewel fish (Coss, 1979), mouse lemurs (Coss,

1978) and mice (Topál & Csányi, 1994) which all show aversive responses to gaze - and gaze cues - such as a reluctance to approach eye cues in a cage or tank.

Cott (1940) noted that animal markings seemed to seek to conceal the eye itself. Cott noted that ‘few natural objects possess greater conspicuousness than the vertebrate eye’, and that it was therefore critical to conceal it if camouflage was to be effective, offering as evidence the image displayed in Figure 1.3. – where the ‘eye’ shape jumps out from other larger and equally high contrast patterns in the picture.

*Figure 1.3. Cott’s diagram demonstrating the inherent conspicuousness of the eye and eye like shapes (Cott, 1940. p. 82)*



The need to hide the eye to avoid detection seems clear, given its highly salient nature – in predators to hide the intent to predate, in prey to avoid detection (Barlow, 1972; Burgess, 1978; Cott, 1957; Neudecker, 1989; Wickler, 1968). And examples of camouflage of the eye abound. There’s the cryptic colouration of the iris in geckos, there are frogs, and many snakes with ‘eye masks’ where the pupil is just a slit and the iris is tinted the same colour as the head. Some creatures such as *Paracheirodon innesi* – better known as the neon tetra (the dichromatic blue/red fish pictured in the eye mask images at Figure 1.4. below) reproduce the pattern of the body through the eye. The short-nosed garfish has a lateral stripe running the length of the body and 'over' the eye.

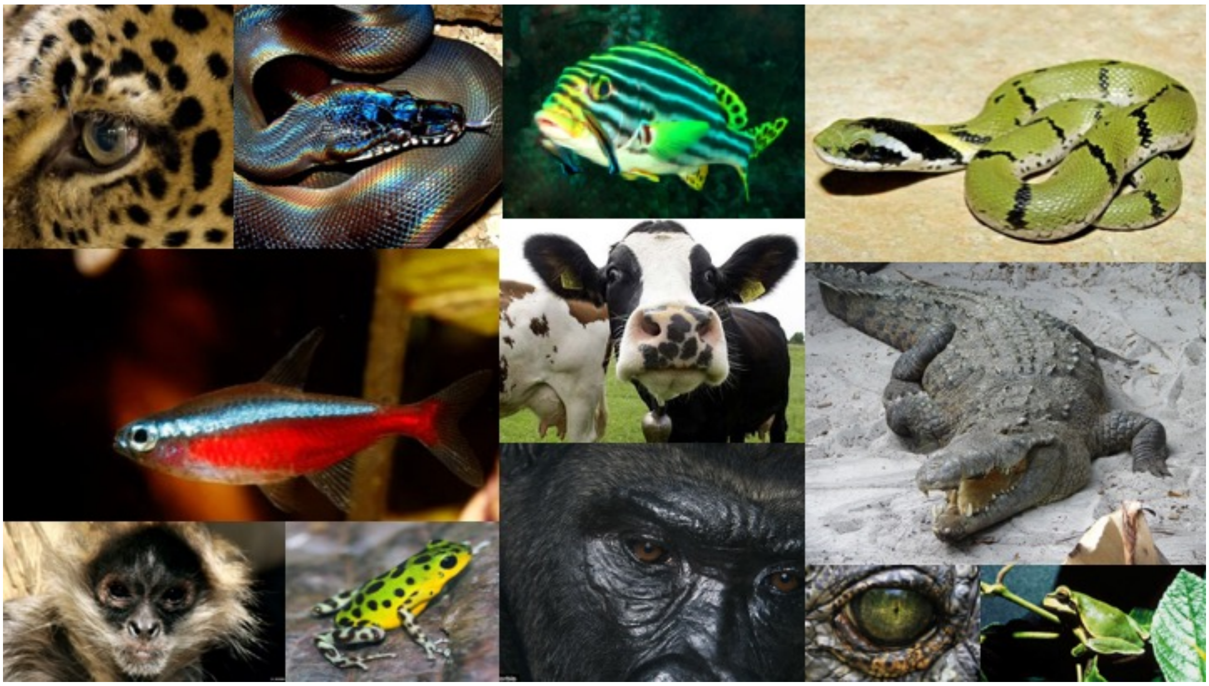


*Figure 1.4. Crypsis: Eye Masks and Disruptive Colouration of the Eyes in Animals.*

It should not be surprising that we pick out the 'eye' in Cott's image so quickly. In nature the eye is ubiquitous, conspicuous, and vulnerable. It has convergently evolved - that is evolved in the same way in multiple different locations and species (Serb & Eernisse, 2008). It's ubiquitous because sight conveys an adaptive advantage (Land, 2014, p. 12). Its basic 'design' – the eye's distinctive shape and dark colouration of the pupil - are dictated by its function (Burghardt & Greene, 1988). The eye's ubiquity and conspicuousness make it a cue to the presence of other animals - predators and prey alike. Thus all species have been under selection pressure to maximise the speed with which they can detect the eyes of others while minimising the chance of their own eyes being detected. In vetebrates the basic design of the eye is not thought to be fundamentally different to the eyes of the earliest jawed fishes – our ancestors some 430 million years ago (Dupret, Sanchez, Goujet, Tafforeau, & Ahlberg, 2014; Land, 2014, p. 28) – consequently, if this is correct, all animals - including humans – have been under evolutionary pressure to detect the eyes of other animals for more than 400 million years. Under such circumstances, perhaps it would be more surprising if animals had not evolved some response to eyes.

Further evidence for the pronounced responsiveness of animals to eye cues comes from the widespread evolution of eye-mimicry – where prey species exploit predators' hypersensitivity to eye cues by presenting eye-spots that can be mistaken for the presence of higher order predators to scare them off, or to fool them into attacking a less vulnerable part of the prey species body (Wickler, 2013; but see Grim, 2013 for an alternative view).

And research has shown just how effective such eye cues can be. Blest (1957), for example, showed that naïve, captive-reared birds were less likely to attack butterflies with eyespots than butterflies that had their eyespots removed. Indeed, some birds were unwilling to attack butterflies with eyespots at all. Blest believed that the combination of eyespot and a startle display – the sudden opening of the wings - might be necessary for the effect to work. To test this hypothesis, he painted eyespots separately from the head on *dead* mealworms – i.e. they could not be more static. Birds concentrated 72% of their pecks to the eyespots rather than the rest of the worm, whereas they directed just 57% at a control image painted on otherwise identical dead mealworms suggesting that when eyespots failed to deter attacks they attracted them. Eyespots might first deter attacks by mimicking predator eyes, and then, when the predator overcomes its hesitancy to attack, the markings subsequently attract the attacks away from more the prey animals more vulnerable areas to minimise damage to the organism. In both examples, the ethological evidence shows that hypersensitivity to eye cues as threat cue can be exploited and cause behavioural outcomes that are not adaptive.

Emery's (2000) review of this ethological evidence is widely cited in the human eye cue literature (2000) [e.g. Grossmann (2017), Burnham and Hare (2007), Bateson et al. (2006) to suggest that humans might have a similar responsiveness to eye cues. While Emery notes a number of studies that found similar effects in primates, Nettle, Cronin, and Bateson

(2013) failed to find an effect in their study examining the effect of eye cues on chimpanzee behaviour. Other studies since Emery's review have continued to demonstrate the effectiveness of eyespots in many animals (e.g. Hossie & Sherratt, 2012; Hossie & Sherratt, 2013; Kjernsmo, Gronholmi, & Merilaita, 2016; Lönnstedt, McCormick, & Chivers, 2013).

The effectiveness of feigning observation has not been lost on humans. In an ingenious programme instigated in 1986, fisherman of the Sundarbans in the Ganges Delta of India were issued 2500 tiger masks to wear on the backs of their heads – a ploy which dramatically reduced the number of attacks from tigers (which prefer to attack unseen from behind) from c.60 a year every year when records began in the 1970s to none in the three years after the intervention (though 29 who did not wear the masks were killed in this period). Unsurprisingly, the fisherman are still wearing the masks today (Safina, 2016). Similar insight has been applied in Botswana to reduce predatory attacks on livestock. One study, for example, reports that lions attacked none of 23 cattle with eyes painted on their rumps, whereas they did kill 3 out of 39 unpainted cows from the same herd during the same period (Safina, 2016).

### 1.2.5. Human Responsiveness to Eyes and Eye Cues

If humans had evolved their own responsiveness to eye-like markings such as is common in so many other species, we would expect to find evidence for this responsiveness in eye tracking studies, cognitive psychological experiments and in neuroscientific research examining how we respond to gaze and gaze cues.

And this is indeed what we find. As in many animal species, eyes draw our attention and provide a key signal of threat (Fox & Damjanovic, 2006). Extensive neuroscientific evidence shows that the processing of facial images is fast, preconscious, shows evidence of automaticity, and cannot be switched off (Palermo & Rhodes, 2007). The eyes are the most important part of the face for conveying non-verbal messages, the first part of the face we look at, the most fixated-on part within the face, and also the most emotionally informative (Haxby, Hoffman, & Gobbini, 2002; Yarbus, 1967). We follow the direction of eye gaze whether we want to or not, even when we are told it will be disadvantageous to the completion of a task to do so (Driver et al., 1999; Friesen & Kingstone, 2003; Kuhn & Kingstone, 2009; Posner, Snyder, & Davidson, 1980).

This attentional responsiveness to gaze begins from very early in pre-verbal infancy, again suggesting it is unconscious and automatic, and it remains a critical social cue throughout our lives (M. H. Johnson et al., 2005) enabling the passing of important social signals which are automatically and effortlessly encoded (Kawai, 2011; Senju & Hasegawa, 2006; von Grünau & Anston, 1995). Eyes are also fixated on even when located outside of the face area. A study by Levy, Foulsham, and Kingstone (2013) showed that it is the eyes on which we fixate even when those eyes are located on the body or hands of a ‘monster’, outside of the face. The conclusion, then, seems clear. Not only are the eyes more important than the face in grabbing our attention, they are also the most important area within the face for communicating information. Eye cues in a poster then, would arguably be especially effective in attracting our attention to the poster’s message, at the very least causing a low-level physical response as our eyes saccade to look at them.

Cognitively, when one of us picks up the direction of another’s gaze and automatically follows it we do so to understand quickly what it is that has captured the other’s attention: more specifically, threat, opportunity, or curiosity (Feng & Zhang, 2014; George & Conty, 2008). From this ‘joint attention’ follows a further automatic cognitive process: perspective taking. Perspective taking enables us to see the world from the perspective of others – and it really does seem to come naturally. Kampis, Parise, Csibra, and

## 1. Introduction.

Kovács (2015) showed that while preverbal infants are unable to sustain object representations when that object is occluded from their own view, they are reflexively able to sustain the object representation if it remains in the gaze of another person that the infant can still see. Thus, they can reflexively ‘perspective take’; they can see the world from the viewpoint of another and sustain an object representation where they would otherwise be unable to. Both joint attention and perspective taking, then, constitute reflexive, automatic and unconscious responses to the gaze of others, and, as such, might be triggered by false cues as well as genuine ones.

Gaze and gaze direction have also been shown to influence our categorization and judgment of others. We are faster to pick out a person’s gender, and we judge their degree of anger, happiness and attractiveness to be greater, if they look directly at us (Adams & Kleck, 2003, 2005; Kampe, Frith, Dolan, & Frith, 2001; Macrae, Hood, Milne, Rowe, & Mason, 2002; Mason, Tatkow, & Macrae, 2005; Mathews, Fox, Yiend, & Calder, 2003; Sander, Grandjean, Kaiser, Wehrle, & Scherer, 2007; Sato, Yoshikawa, Kochiyama, & Matsumura, 2004; Vuilleumier, George, Lister, Armony, & Driver, 2005). Thus, gaze can also affect our cognitive processing of information, suggesting that eye cues too might influence our categorization and judgement of a situation.

Our physical response to gaze, often triggered in the aforementioned experiments not by real eyes but by photographs of eyes, is to look at them and then look where they are looking, cueing the cognitive processing of the social signal – is the watcher angry, happy, sad, etc. - and enabling us to see things from their perspective. Photographs of eyes are eye cues, indeed cropped photographs showing only the ‘strip’ of the eyes are frequently used in eye cue experiments. Thus, eye cues might affect our decision-making via the same cognitive mechanism: cueing us to see ourselves, and/or the situation we are in, through the ‘eyes’ of another, judge ourselves or the situation from a second person perspective, and leading us, perhaps, to adjust our behaviour.

### 1.2.6. Neuropsychological Research

Our neural response to being watched is also fast and automatic. Despite what would seem the obvious evolutionary advantage of following the gaze of another to see where they are looking – to detect predators, prey, opportunity or threat – and despite the fact that we do follow gaze automatically, unconsciously and rapidly, we nevertheless prioritize the emotional content of gaze ahead of its direction (Conty, N'Diaye, Tijus, & George, 2007). Within 150-170ms of detecting gaze our brains have already begun to form an impression of the intentionality behind it (Conty et al., 2007) and recruited elements of the ‘social brain network’ to begin to form impressions and subconsciously shape our reactions (George & Conty, 2008; M. H. Johnson et al., 2005; Senju & Johnson, 2009).

In contrast, we don’t begin processing of gaze direction until after 190ms. This is surprising: to have evolved to first process the emotional content of gaze, to ‘read the mind in the eyes’ (Baron-Cohen, 1997a; Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001) ahead of following gaze direction suggests that knowing and responding to the emotional content of the gaze of others, rapidly and automatically, has been more important for our survival than looking at where others are looking – more important than seeing predators or prey.

This neural sensitivity to gaze and rapid response to it comes from the >30 regions of the brain dedicated to visual processing (Emery, 2000). The neural response to eyes is routed via a subcortical face detection pathway including the superior colliculus, pulvinar and amygdala (Senju & Johnson, 2009). This circuit appears to attach emotional significance to the eyes (such as threat or embarrassment) and then sends signals to the brainstem to initiate an emotional response (such as an increased GSR or blushing; (Emery, 2000)). Emery (2000) suggests that this could be achieved ‘via a direct projection from the central nucleus of the amygdala to the brainstem’. Thus, the effect of being watched would not only be rapid, but automatic and below conscious awareness.

This automatic, rapid 150-170ms response enables us, as mentioned above, to infer the emotional states of others, to ‘read their minds in their eyes’, and it seems to do so in order to inform action (Baron-Cohen, 1997a; Baron-Cohen et al., 2001; Conty, Dezecache, Hugueville, & Grezes, 2012; Schilbach, 2015). This is shown by the way in which our perception, and ultimately the neural processes responsible for action, are affected by being watched.

We perceive objects differently, for example, when they are located in the direction of another’s gaze. To illustrate, Bayliss, Paul, Cannon, and Tipper (2006) showed how our affective judgement of objects can be modulated by gaze cues. When these authors presented participants with objects in the presence of a face image looking towards that object, participants rated the object as more likeable than in a presentation when the face was looking away from the object. However, when arrows instead of faces were used to point at or away from an object, directional cueing of eye gaze was not accompanied by modulations in affective response. This suggests that the presence of an observer’s gaze has a real effect on how we perceive and respond affectively to an object (Bayliss et al., 2006; Hartendorp et al., 2013; Manera, Elena, Bayliss, & Becchio, 2014).

Event Related Potentials research – the use of Electroencephalogram (EEG) scanning of neuronal electrical processing during the presentation of specific stimuli - constitutes another fruitful line of inquiry. Studies have shown, for example, that objects that are gazed at by another are processed by infants as familiar, whereas objects that are not observed by another are processed as unfamiliar and novel – and, moreover, that when it comes to infants themselves, being watched actually has a direct effect on the neural circuitry they use to perceive and process objects (Becchio, Bertone, & Castiello, 2008; Nelson, 1997). In one such experiment, for example, infants watched a video of an adult gazing towards one of two objects. Then a paired-preference test phase began: the same objects were presented but without an adult gazing at them. Results showed that infants looked reliably less at the previously ‘gazed’ object, suggesting that the object cued by the adult’s gaze was perceived as less novel than the non-cued object. (Becchio et al., 2008; Manera et al., 2014) Moreover, a replication of this experiment in which the infants brain activity was measured with EEG showed that the object was processed via different neural networks if it had been looked at by an adult previously (V. M. Reid, Striano, Kaufman, & Johnson, 2004).

Further investigation involving electromyography (EMG) shows that direct gaze stimulates emotional arousal in the brain (Conty et al., 2010), while positron emission tomography (PET) scanning reveals that direct gaze activates specific areas of the brain such as the right amygdala that averted gaze does not (Kawashima et al., 1999). In their review of direct gaze effects, Becchio et al. (2008) sum up the picture thus: “Under the gaze of others the object is enriched with motor, emotive and status components that it would not display if not looked at”. In other words, under direct gaze we process information via different neural pathways and perceive and respond to objects differently.

Most significantly of all, perhaps, a separate line of research suggests that we use different neural circuitry to control our actions when we are watched. A series of fMRI studies examining differences in performance in simple stimulus-response compatibility (SRC) tasks have shown that when we are watched performing simple tasks (e.g. button pressing) neural activity occurs principally in the inferotemporal cortex, the amygdala and medial prefrontal cortex, whereas tasks performed when we are not being watched leads to principal activation of the extrastriate visual and posterior parietal cortices (Schilbach et al., 2011). This suggests that the neural circuitry responsible for social perception plays a greater role in the generation of motor-control when we are watched than when we are not. Indeed, Schilbach et al. conclude that being watched ‘significantly changes the neural underpinnings of action control and recruits brain regions previously implicated in action monitoring, the reorienting of attention and social cognition’.

Similarly, fMRI studies of mimicry under conditions of being watched and not watched suggests that gaze causes activation in an area of the brain on the boundary between the posterior and anterior of the medial Pre-Frontal Cortex (mPFC). The posterior region of the mPFC is associated with response inhibition, while the anterior is associated with mentalising and social cognition, thus our neural response to gaze would seem to inform motor action in part by aiding us in inhibiting responses while also helping us to understand the emotions of others (Yin Wang & Hamilton, 2014; Y Wang, Ramsey, & Hamilton, 2011).

Recent evidence from animal research confirms similar neural mechanisms operating in monkeys and identifies two relevant and corresponding populations of neurons in primates: ‘social neurons’, which are preferentially active when a conspecific is present, and ‘asocial neurons’ that are preferentially active when the monkey is alone (Demolliens, Isbaine, Takerkart, Huguet, & Boussaoud, 2017). The key finding here is that when social neurons are active monkeys perform much better on a visuomotor task. Since the activity of social or asocial neurons depends only on the presence or absence of a conspecific the suggestion is thus that in some primates at least there is a pronounced responsiveness to the presence of others that strongly resembles our own (Demolliens et al., 2017).

### 1.2.7. Evolutionary Evidence Part 2: Evolutionary Psychology

In summary, the available evidence to date shows that humans, like many animals, are hypersensitive to the presence of eyes. Indeed Baron-Cohen posits that humans have a neural module he describes as the Eye Detection Device (EDD) that is hypersensitive to the presence of eyes, and a module he calls the Intentionality Detector (ID (Baron-Cohen, 1997b, pp. 31-38) which detects the presence of other humans or animals. The concept of the ID comes from the human tendency towards apophenia, defined by Brugger (2001) as a 'pervasive tendency of human beings to see order in random configurations'. We know this best when people talk of seeing faces in the clouds, a man in the moon – or from news stories where individuals report seeing Jesus in toast, or the Virgin Mary in a tortilla (Liu et al., 2014). Conrad (1958) describes apophenia as the 'unmotivated seeing of connections [accompanied by] a specific feeling of abnormal meaningfulness' – a tendency, in other words, to perceive patterns in the random, to commit perceptual 'type 1' errors, to see 'something' when nothing is present. Viewed through the lens of evolutionary psychology this is a necessary flaw – far better to exert unnecessary caution in response to that rustle in the bushes than to make a balanced judgement that it probably isn't a predator and be eaten. Baron Cohen's theories neatly tie together the insights we have gained in our review of the ethological and cognitive neuropsychological evidence. He suggests the ID & EDD work together to trigger a Shared Attention Mechanism (SAM), which enables us to perspective take, and a Theory of Mind Mechanism (TOMM) that enables us to automatically adjust our behaviour to our audience (Baron-Cohen, 1997b, pp. 31-58).

Thus, in Baron-Cohen's theory, eye detection and agency detection are hyper-sensitive, and adaptive. The eye-tracking evidence for his EDD – some of which is outlined above and shows how we pay attention to eyes, look where they are looking etc. - is extensive (Baron-Cohen, 1997b, pp. 31-50), the evidence for an intentionality detector is sparser, but still persuasive. Baron Cohen cites studies showing how easily we anthropomorphise shapes (Heider & Simmel, 1944) and evidence from individuals with focal brain damage that has enabled neuroscientists to identify localised neurological modules responsible for distinguishing objects from agents (Warrington & Shallice, 1984). The neurological evidence for perspective taking – again as presented briefly above - supports the idea of a Shared Attention Mechanism. What Baron-Cohen provides us with is a useful four-stage model: EDD-ID-SAM-TOMM, to summarise how many of the separate lines of evidence might operate together.

More recently, Conty, George, and Hietanen (2016) have presented an updated and largely complementary, if slightly more conservative, explanation of the mechanism by which direct gaze might influence behaviour. As with Baron-Cohen's EDD, they argue that gaze's highly salient nature first captures attention. They suggest the main secondary effect is similar to Baron-Cohen's Shared Attention Mechanism heightening self-referential processing. In the last step of their model gaze prompts heightened concern as to how we are socially evaluated, thus prompting reputational concerns and leading us to moderate our behaviour (Conty et al., 2016). Thus Conty et al.'s model is largely aligned with Baron-Cohen's. Importantly for our discussion, they introduce the idea that eye cues might prompt reputational concerns and suggest that the human sensitivity to gaze might be triggered by eye cues too.

From the evidence presented, the importance of detecting and responding to eyes in humans is clearly not limited to predator-prey detection. If it were, it seems probable the human eye would have evolved to be less salient to avoid attention – as in the many animals with various ways of camouflaging their eyes described previously.

### 1.2.8. Eye Morphology: Can we Read the Mind in the Eyes?

In fact, biologically, our eyes may have evolved to do the opposite, to capture attention, and to communicate. We are the only primate with such a large, pronounced white scelra and the only animal in which the colour of the sclera can change - redden or yellow - to communicate vitality, emotion or illness (Provine, 2013; Provine, Cabrera, & Nave-Blodgett, 2013). Biologists have suggested that humans evolved this ‘design’ in order to give us ‘signalling’ eyes to enable our hyper-sociality as opposed to the camouflaged eyes of other species (H. Kobayashi & Kohshima, 1997, 2001, 2011; Provine et al., 2013).

This saliency of the human eye is suggestive of their greater import in human evolution, and perhaps intimates that our sensitivity to eyes and gaze should be greater than that of most animals. It certainly makes intuitive sense that our cooperativeness and degree of inter-dependence may have made having eyes that seek to capture attention, rather than avoid it, evolutionarily advantageous, in precisely the same way that in other animals the cost of having more visible eyes would likely have been too great. Indeed, this latter suggestion provides a possible explanation as to why so many creatures have actually evolved the capacity to hide them.

But we can do much better than mere intuition. Research testing the effects of human eye morphology have found it plays an important role in communication (H. Kobayashi & Kohshima, 1997, 2011; Provine, 2013). For example, in one study just increasing the redness of the sclera in pictures of human eyes that were otherwise identical was enough to lead participants to describe those individuals whose eyes were displayed as less happy, and likely experiencing greater anger, fear, sadness and disgust (Provine, 2013).

That the eye’s morphology evolved not only to allow us to see out but also to allow others to see in - ‘reading the mind in the eyes’ as Baron-Cohen put it (Baron-Cohen, 1997a; Baron-Cohen et al., 2001) - is increasingly well evidenced. One study has shown, for example, that the eyes may play as important a role in human courtship and seduction than any other non-verbal display, and are perhaps even the most important non-verbal factor (Givens, 1978). Further evidence that the eyes communicate is shown in a study that demonstrated that increased pupil diameter was correlated with feelings of lower self-esteem in females (Zeitner & Weight, 1979; Hess, 1975). Others have demonstrated that gaze direction alone can communicate a range of different emotions (E. Herrmann, Call, Hernandez-Lloreda, Hare, & Tomasello, 2007; Tomasello, Hare, Lehmann, & Call, 2007). Finally, while many animals produce 'tears' as oculomotor lubrication to remove grit and disease, it might be that the communicative role of tears as an emotional signal is unique to humans (Provine, Krosnowski, & Brocato, 2009). The language of the eyes can communicate a surprisingly wide range of emotions.

Computer scientists have shown that our emotional reactions to images can be classified with 90% accuracy based on eye-movement alone (Raudonis, Dervinis, Vilkauskas, Paulauskaite-Taraseviciene, & Kersulyte-Raudone, 2013) and that eye tracking can detect when we are disgusted, ashamed, or experiencing pleasure at similar levels of accuracy (Maskeliunas & Raudonis, 2016). If this seems remarkably strong evidence for the role of eyes in communicating emotion, a recent computer science study suggests that soon we might be able to prove that you can 'read the mind in the eyes' in greater detail than anyone would ever have imagined. Computer scientists at Flinders University in Adelaide asked study participants to take a 10 minute walk from a computer science laboratory to a shop and back while wearing state-of-the art eye tracking glasses. The researchers were able to show that they could use machine learning algorithms to predict participants' levels of curiousity in two questionnaires, as well as their scores on the 'Big Five' measure of personality, at rates well above chance *from eye movement alone* (Hoppe, Loetscher, Morey, & Bulling, 2018). If eyes can communicate not just our emotions but who we are at such a deep level it becomes difficult to argue that we did not evolve their morphology and movement in part to communicate.

In addition to the strong evidence from psychology demonstrating that our eyes evolved to signal emotions and the persuasive evidence emerging from computer science to the same end there is also evidence to suggest that eyes played a critical role in the evolution of sociality.

Building on the evidence that there really might be a language of the eyes, Kobayashi, Ohtsuki, and Wakano (2016) found that increases in eye features shown to have a communicative function - the exposed sclera and increased horizontal elongation in eye morphology - correlated with increased neocortical size, and group size, across 30 primate species, including humans. Similarly, in a separate study, they found that the extent to which the eyeball was able to move when scanning predicted neo-cortex size and group size. In what they describe as the *gaze grooming hypothesis,* (to which we will return in our more detailed discussion of evolutionary psychology and surveillance below) Kobayashi and Kohshima make an important contribution to growing evidence for the role of the eyes in developing humans as a functional social creature, suggesting that eyes allowed us to communicate, reassure, reproach, threaten and seduce pre-linguistically. These social mechanisms likely built, and capitalized, on pre-existing attentional hyper-sensitivity to eye regions. Our survival and fitness arguably depended more on our ability to respond to the messages encoded and transmitted by the eyes of our conspecifics than on hiding our eyes to avoid detection by our predators or our prey. Given, then, that we have evolved to both send and receive information via gaze, eye cues – just like eyes themselves – might similarly capture our attention and communicate messages.

At this point, let us briefly pause to recap on what we have learned so far. On the one hand, the evidence shows that we are hypersensitive to gaze, it shows how important the gaze of others is for human cooperation, and that our eyes might have evolved not just to see out but to communicate. On the other, social surveillance changes the neural networks by which we process information, changes how we perceive objects and might change behaviour by cueing us to consider the reputational effects of our actions, by moderating anti-social urges, and by promoting prosocial actions.

To understand this is to begin to speak in the language of evolutionary psychology as applied to the evolution of cooperation, where avoiding anti-social behaviour and taking prosocial actions is done in the hope of securing longer-term rewards through direct and indirect reciprocity – that is, through increased cooperative benefits derived from being *seen* as reliable, trustworthy and astute (Axelrod & Hamilton, 1981; Trivers, 1971). Although debate continues as to whether cooperation first evolved through direct or indirect reciprocity, or through costly signalling (Tanaka, Ohtsuki, & Ohtsubo, 2016), each depends on individuals managing the impression they create on others – signalling their trustworthiness, in other words - before the gaze of the group.

This unique combination of social and evolutionary pressures on the human response to eye gaze raises an intriguing possibility. Under such circumstances human group size would arguably have been limited by the number of reputations that individuals could remember – we couldn't afford to admit anyone whom we didn't recognise and whose reputation we could not recall (Humphrey, 1976). Any expansion of the group beyond this constraint would risk free-riding, allowing cheaters to prosper and risking the breakdown of cooperation.

Dunbar's *social brain hypothesis* holds that human groups *were* limited in size, to around 150 in fact, constrained by neocortical limits on the ability to recall and manage the relationships and reputations of others (Dunbar, 1992, 1998). Reputations would also have been built on what we saw others doing, and on third party observation and report. To cheat and be caught would have resulted in punishment, in turn enhancing cooperation (Fehr & Gachter, 2002). Since failing to punish can quickly result in the widespread breakdown of cooperation (Fehr & Fischbacher, 2003), and conformity is critical to group survival, to punish those that defect from rules and norms is essential (Cosmides, 1989; Tomasello, 2009; Tomasello & Vaish, 2013). To be hyper-sensitive to eyes, then, would have been adaptive – avoiding potential damage to one's reputation which might have had costly consequences.

Bottom line? Early societies would surely have needed to live in a primeval version of the panopticon. And they did. Cosmides and Tooby (1997), for example, have described how "*…each of our ancestors was, in effect, on a camping trip that lasted an entire lifetime, and this way of life endured for most of the last 10 million years"* – just like on a camping trip today there was no place for free-riders – everyone saw how much you did or didn't contribute. Back in the day of our prehistoric ancestors, almost everything people did would have been observed by other group members.

For those that *did* breach the norms and rules there was little chance of getting away with it unseen, and any transgression would have damaged their reputation amongst the collective as word spread. In this regard, evolutionary notions concerning the origins of gossip have interesting implications for theories of surveillance. Gossip, so the argument runs, allowed the eyes of others to 'watch' when yours weren't there (Dunbar, 2004). It provided a shared, group-level view of people's behaviours. Let us return, in this context, to Kobayashi, Ohtsuki, and Wakano's (2016) *gaze grooming hypothesis,* based, as previously described, on evidence that human group size correlated with the extent to which primates (including humans) have signalling eyes. Kobayashi et al.'s work suggests that 'gossip', in the form of disapproval, anger or disgust might have been initially, pre-linguistically, communicated by the eyes. However, as group size increased it seems unlikely that this would have been sufficient to communicate and share the reputational information of people who one rarely saw, and so this inference integrates seamlessly with the arguments in evolutionary psychology which suggest that it was the need for gossip to regulate increasingly large groups that drove the evolution of language (Dunbar, 1993, 1996). The language of the eyes, rich and complex though it was, exceeded its functional limit under the pressure of our increasing group size and more complex relationships - requiring *homo sapiens* to evolve a verbal language to keep cooperation functioning. This evolution of a verbal language enabled us to hear gossip about things we didn't immediately observe. With more complex verbal language we could adjust our 'view' of people to include everything they did, enabling us to track people's reputations and so to live in larger groups (Dunbar, 2004; 2014, pp. 264-272). We might not have been able to watch everyone, but we didn't need to. Everyone was watching each other as part of the collective. And although communication from the eyes was now augmented by verbal language, the evidence previously presented suggests that the language of the eyes still remained rich and informative.

Gossip, then, in word or glance, was (and still is) used to select cooperative individuals to interact with, to ostracise the selfish, and to mitigate egoistic behaviour by ensuring that the least egotistical, least selfish individuals were the most prestigious and thus most likely to be chosen by females as mates, and others as friends, in activities for mutual benefit (Baumeister, Zhang, & Vohs, 2004; Feinberg, Willer, & Schultz, 2014; Henrich & Gil-White, 2001). Indeed, in this regard the power of ostracism cannot be overstated: research shows that its effects on the human brain are as keen, perhaps even keener, than those of physical pain (Eisenberger, Lieberman, & Williams, 2003; MacDonald & Leary, 2005; D. S. Wilson, Wilczynski, Wells, & Weiser, 2000).

Although it seems probable that exclusion alone would often have been enough to have motivated our ancient ancestors to drop their selfish tendencies and cooperate, one can readily imagine what would have happened to selfish individuals who persisted in their anti-social behaviour. The uncooperative might have not just been ostracized but physically or materially punished into the bargain, reducing their status and thereby their attractiveness to other members of the group, and in turn, their chances of reproducing. At the extreme, they might have been expelled from the group altogether, thus becoming much more vulnerable to predators and famine - and as a last resort even executed, to rid the tribe of their parasitic behaviours and to serve as an example to others (Black, 2014).

To develop our understanding of how our sensitivity to the language of the eyes in our ancestral environment might still affect our behaviour today, the concepts of ultimate and proximate mechanisms in evolutionary psychology are critical (Mayr, 1963; Tinbergen, 1963). Ultimate explanations are concerned with why a trait or behaviour exists, that is why natural selection has favoured its existence through explaining how it contributed to enhancing an organism's adaptive fitness. Proximate explanations explain how it manifests itself (Scott-Phillips, Dickins, & West, 2011). The idea is perhaps best explained by Krasnow & Delton (2016) who argue that behaviour can have a deep evolutionary logic underlying it, but that this same logic can manifest itself in responses and drives that can be irrational when operating in new environments. For example, deep evolutionary logic gave us a rational drive to maximise calorific consumption in our ancestral environment when lives were active and food could be scarce. However, the proximate mechanism is our love for dense, fatty, sugary, high calorie foods – irrational in the modern context of sedentary work, food surpluses and obesity. The proximate mechanisms can misfire while the deep evolutionary logic remains sound. In our discussion of ethology birds avoiding the ocelli on butterfly wings for fear they may have caught a glimpse of a higher-order predator is a misfiring of a proximate mechanism shaped by the deeper evolutionary logic of predator avoidance (De Bona, Valkonen, López-Sepulcre, & Mappes, 2015). Whenever the costs of decision errors are asymmetrical, evolution should select for psychological mechanisms that avoid making a costly error at the expense of making lots of cheap errors (Krasnow & Delton, 2016). It doesn't get much more asymmetrical than being ostracized, punished, having your chances of reproducing damaged, being expelled, starving, being eaten or even being executed because you didn’t pay attention to the eyes and gaze of others. Consequently, being acutely sensitive to the presence of eyes would be adaptive even if it led to one making occasional errors by assuming you were watched when you weren’t.

This is perhaps why in agent-based simulations of human interactions, natural selection favours agents willing to cooperate even with those they’ll never see again (Delton, Krasnow, Cosmides, & Tooby, 2011). On a life-long camping trip the ultimate cause - an acute sensitivity to social gaze would have constituted a pronounced evolutionary advantage. Correspondingly, therefore, the proximate psychology might then demonstrate a bias towards gaze detection and the avoidance of errors in situations where whether one is watched is uncertain, since defection errors are costlier to one’s survival and fitness than unnecessary cooperation.

### 1.2.9. The Supernatural Surveillance Hypothesis

The camping trip began to end after c.7000 B.C when humans first pioneered a new way of life in agricultural communities, food surpluses increased group size (Bar-Yosef, 1998) and agriculture demanded more manpower creating rapid population growth (Barker, 2006, pp. 32-33, 399-400). Now we could hide in the crowd, disappear into the anonymity of large towns and eventually vanish into cities or spend our lives moving between groups across extended trading networks (Nock, 1993, pp. 1-16). If we evolved to live always in the spotlight of social gaze, the agricultural revolution presented a threat: we were escaping the primeval panopticon that had regulated our behaviour and allowed cooperation to evolve. The way in which we responded to this suggests that we might be so sensitive to gaze that we are able to adapt to ‘feel watched’ even when we are far from the gaze of the group.

Evidence for the rapid growth of population with the founding of the first permanent settlements is found in the archaeological record. To take one of the most notable examples,

## 1. Introduction.

in c.13,000 BC the oasis at Jericho in Palestine supported one nomadic hunter-gatherer tribe of c.100 people (Harari, 2015, pp. 87-94), living in more or less the way anatomically modern humans had for 300,000 years (Dunbar, 2014; Strait, Grine, & Fleagle, 2015). By around 8500 BC, some 500 years later, the first permanently settled agricultural villages had been founded, the cultivation of wheat and barley had begun, rapid population growth was underway and there is evidence of some humans living in small cramped villages of around 1000 (Harari, 2015, pp. 87-94). Within a thousand years, not so far away, the population at Catalhoyuk a town founded in c.7500 BC in Anatolia, is estimated to have housed on average around 5000-7000 people, up to a maximum of 10,000 (Maynes & Waltner, 2012, p. 12). In 1500 years, not much longer than the blink of an eye in evolutionary terms, some human groups had gone from living in bands of c.150 to groups of up to 10,000.

As groups grew in size, exceeding 150 people, they outgrew our cognitive capacity to keep up with the latest gossip and track the reputations of all group members. Cooperation based on reputation tends to break down in larger groups (Boyd & Richerson, 1988; Dunbar, 1992; Panchanathan & Boyd, 2003). Consequently, larger groups will tend to collapse, fission or feud unless they develop new ways to expand the circle of trust. Trust requires an extension of surveillance: trust but verify the Russian proverb tells us. But how to trust when you cannot verify?

We needed a way to bring people back into the spotlight of social gaze, and the light of religion provided the answer. For perhaps as long as 500,000 years prior to the Neolithic God was experienced primarily through ritual and mysticism, and religion was played out through theatrical acts, binding the group, sustaining hierarchy, making room for new members (Dunbar, 2017). Dunbar suggests that over the Neolithic period we scaled up the ritual that had bonded our groups of c.100-c.150 so as to create much larger groups. In doing so we repurposed our higher order theory of mind mechanisms to imagine Gods capable of threatening future punishment in order to sustain the norms and social hierarchies that held larger groups together (Dunbar, 2013). We evolved to feel watched by Gods to keep us in order.

First, less us consider the circumstantial evidence for this. The emergence of ‘big Gods’ coincides with emergence of agriculture (Swanson, 1964, p. 77). In South-Eastern Turkey, in the region where the agricultural revolution began, we find in the archaeological remains of Gobekli Tepe (Figure 1.5.) the first monumental pillared structures ever produced. These date from 9500 BC just as the agricultural revolution was getting started and are just 30kms from where geneticists place the origins of one of the world’s first domesticated wheat variants (Harari, 2015). Similarly, the first megalithic architecture in the region of Ireland, Britain, Denmark, and southern Sweden, the most famous being Stonehenge, dates from c. 4000-3900 BC (Hutton, 1993), the same time as the beginning of agricultural revolution in these countries (RowleyConwy, 2004). Both Gobekli Tepe and Stonehenge were vast burial sites, containing the remains of thousands of the dead over many centuries (Darvill, Marshall, Pearson, & Wainwright, 2012; Willis et al., 2016). Agricultural revolutions often arrived with ancestor worship (Armstrong, 2004, pp. 43-63). The vast burial sites, and monumental architecture may have been playing a role in extending surveillance so you couldn’t escape gaze and judgement: the dead were watching you wherever you went.

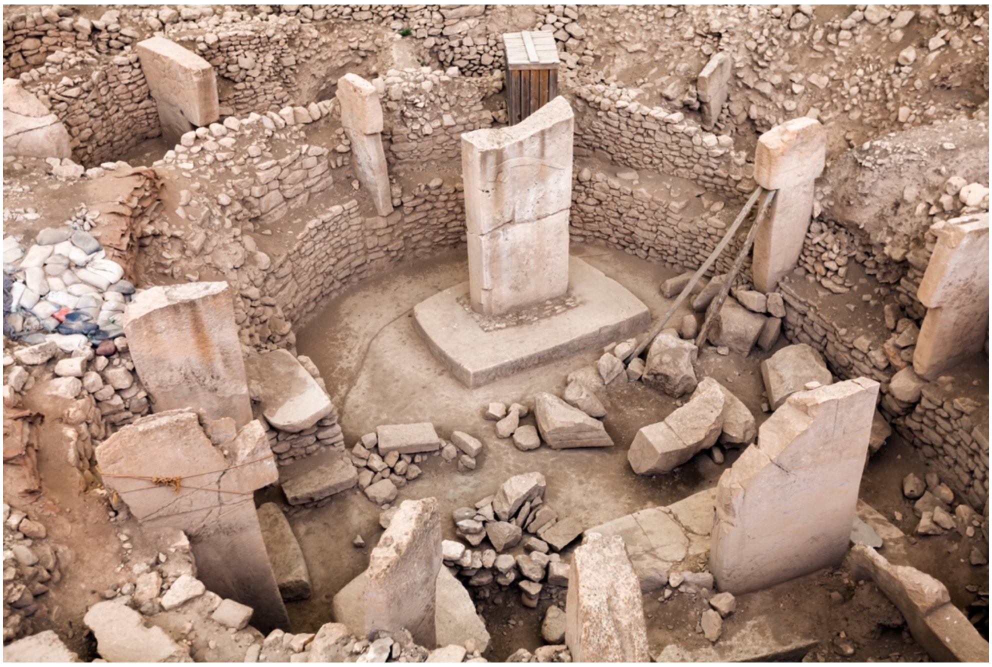


*Figure 1.5. The remains of the temple at Gobekli Tepe ("Potbelly Hill") in modern South-East Turkey. Built c.9500 BCE the same time as the agricultural revolution began, or perhaps even just before, the first monuments were circles of massive T-stone pillars – megaliths built staring down across the surrounding countryside.*

Hawaii is a powerful case study, showing how important 'feeling watched' has been to sustaining humans in larger groups. Hawaii's agricultural revolution began in the mid-15th Century bringing with it an ancestral spirit-based faith and standing stones – *Heiau* - used to mark field boundaries and cultivated fishponds (Mulrooney & Ladefoged, 2005). *Heiau* were small-scale family shrines erected to the dead by the growing kinship group. They were said to be inhabited by local spirits responsible for assuring the productivity of nature: ever-present, all-seeing. They demanded that the group avoid polluting the environment by keeping a number of religious taboos (Mulrooney & Ladefoged, 2005).

The *Heiau* provide evidence that Hawaii evolved religion to help keep groups functioning at an ever-larger scale through the island's agricultural revolution. They were so placed as to overlook the landscape (Mulrooney & Ladefoged, 2005), to ensure the ancestor-gods were always watching you and rules were obeyed. So, too Gobekli Tepe and Stonehenge were built with sweeping vistas (Curry, 2008; Hutton, 1993; Willis et al., 2016).

Religious monuments gave tangible existence to spiritual ideas, made them more immediate, reinforced the belief structures, ‘nudged’ them to behave better by reminding everyone that God was watching. Building these early ‘cathedrals’ on a hill only amplified their function (for a Buddhist example see Figure 1.6.).

This surveillance didn’t supplant social surveillance, but extended it in a much more literal way than we might intuitively perceive. Gobekli Tepe, Stonehenge and *Heiau* were all temples to the dead, reminding people of the dead’s continued existence (Curry, 2008; Hodder, 2010; Hutton, 1993; Willis et al., 2016). This was not an other-worldly existence, of heaven or hell as we might conceive it. Rather the dead remained very much present, became ancestor-gods and lived among the tribe. They gained new powers, could be in multiple places at once, knew your thoughts and watched everything you did. You were always watched. Even today Hawaiians demand to know which spirits live in a house before buying it, and do not speak ill of the dead who are, as they conceive it, in the room and quite capable of offence (D. Johnson, 2016, p. 54).

All over the world, as the agricultural revolution took hold, as group size exceeded human ability to watch, track, and remember the reputations of everyone they interacted with, groups developed a belief in the constant physical presence of the dead. Archaeologists studying early agricultural settlements at 51 sites across the Near East looked for evidence of storage facilities for surplus harvest as a proxy indicator for groups that gained the capacity to grow (Rodrigue, 1992). Those groups that showed such evidence of growth consistently developed religion shortly after they developed surpluses (Rodrigue, 1992). Similarly, anthropologists have shown in a study of 168 societies that small hunter-gatherer communities rarely believe in interventionist, all-seeing gods, that larger social groups usually do, and that agricultural and pastoral communities are significantly more likely to believe in such omniscient, omnipresent gods than hunter-gatherers (Peoples & Marlowe, 2012). This is further supported by analysis of data from the *Standard Cross-Cultural Sample (SCCS)* database, which tracks over 1800 carefully controlled variables for the 186 societies it describes, which showed that across those 186 human societies belief in watchful, moralizing Gods consistently correlates with larger group size and higher group cohesion (Roes & Raymond, 2003), and indeed, according to renowned 20th Century anthropologist Leslie White *"all the great religions of the world today, with the possible exception of Islam, may be regarded as the products of the Agricultural Revolution. Confucianism, Taoism, Buddhism. Hinduism, Zoroastrianism, and its offshoot Mithraism…"* (L. A. White, 1959, p. 362).

As societies grew larger, religion helped to bind them together beneath the gaze of ancestors, gods, or God, with their associated religious rituals allowing social groups to remain cohesive while reducing the need for physical and social intimacy and physical proximity (Watson-Jones & Legare, 2016). The gods were watching you, tracking your behaviour.

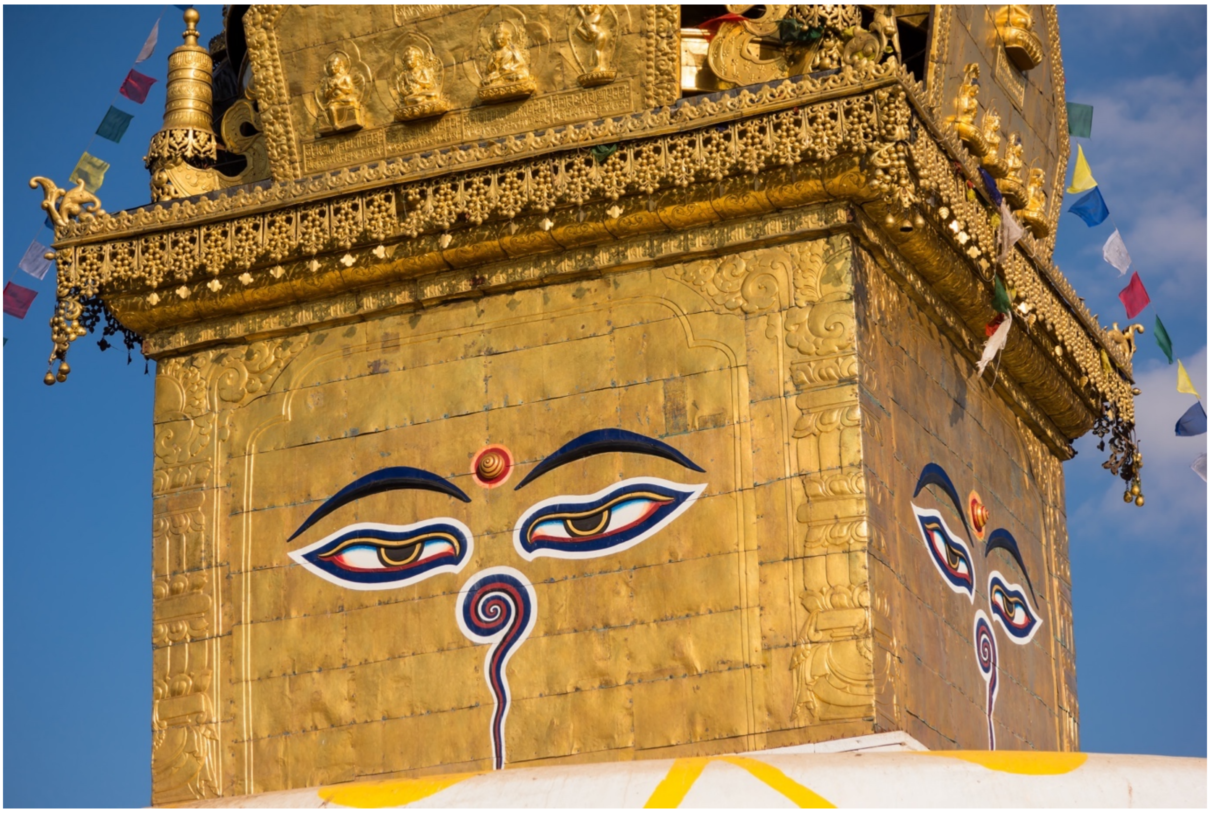


*Figure 1.6. The 'wisdom eyes' of Buddha in Swayambhunath Stupa, Nepal stare down from the tower on top of the hill in all directions – with an inescapable all-seeing third eye between the brows.*

Beyond the archaeological and anthropological record is more direct evidence that humans have often felt watched by Gods and spirits. Psychologists have suggested that we may have a cognitive system dedicated to forming illusory representations of other intelligent agents (Jesse M Bering, 2006; Jesse M Bering, McLeod, & Shackelford, 2005; Jesse M. Bering & Shackelford, 2004) and that the neural networks involved in thinking about God are the same as those that are active when we are watched by others (Boyer, 2003; Schjoedt, Stødkilde-Jørgensen, Geertz, & Roepstorff, 2009). All of this leads to the *Supernatural Surveillance hypothesis* which posits that natural selection might have favoured those with a greater neuropsychological disposition towards belief in God to overcome the selfish interests that would otherwise have seen trust break down and groups fragment once they grew larger than the 150 reputations we could track (D. Johnson & Bering, 2006).

## 1. Introduction.

The most direct evidence for the supernatural surveillance hypothesis comes from economic games run in eight diverse communities from Vanuatu via Siberia, to Tanzania and Brazil (Purzycki et al., 2016, 2018). Purzycki et al. (2016) found that the more omniscient, omnipresent and punitive participants rated their moralistic gods to be, the higher the proportion of coins they allocated to geographically distant co-religionist strangers rather than to themselves or co-religionist neighbours. In other words, before the eyes of God, people were less selfish and more cooperative.

In laboratory studies the feeling of being watched has also been induced indirectly to reduce cheating. An increase in pro-social behaviour was found in children who thought an imaginary friend was present when playing games (Piazza, Bering, & Ingram, 2011), in adults when the 10 Commandments were displayed (invoking the idea of an all-seeing God) (Piazza, 2012), and in lab students when they believed the ghost of a former student may have been present (Shariff & Norenzayan 2007). Most pertinently, Will Gervais and Ara Norenzayan in their study *Like a Camera in the Sky? Thinking about God increases public self-awareness and socially desirable responding* (2012) showed that the religious respond to being primed with thoughts of God in the same way as they do to being primed with thoughts of others watching them: by becoming more self-aware and self-conscious.

In summary, the supernatural surveillance hypothesis suggests that the feeling of being watched is not necessarily dependent on eye morphology, and might transfer to more abstract forms of surveillance. Certainly by Gods, possibly by other agentive apparitions, and possibly, therefore, by Foucault's abstract *technologies of control* - by cameras, and apps, and all the components of the surveillant assemblage.

### 1.2.10. Evolutionary Evidence Part 3: Cheating in Darkness

In making the case for why there might be a watching eyes and perceived surveillance effect our final piece of evolutionary evidence is that which shows that feeling less watched make us more inclined to cheat. Zhong, Bohns, and Gino (2010) showed that making a room slightly darker led participants to cheat more, even when they were clearly visible. They suggest that darkness was a false cue, making participants feel more anonymous even when they weren't. A follow up study similarly revealed that participants who were watched cheated more when looking through the darkened lenses of sunglasses than through clear lens glasses. Again Zhong et al. (2010) suggest that it was the perceived anonymity induced by the perceptual sense that they were in the shade and shadows that led to greater cheating, despite the fact that participants were equally visible to the experimenter in both studies. This laboratory evidence suggests that humans have evolved to respond to unconscious cues to their own visibility even when conscious consideration of their situation would not justify any behavioural change.

Research on the presence of street lighting on crime suggests that the influence of unconsciously perceived visibility on behaviours may transfer beyond the laboratory. In a meta-analysis of 14 studies of street lighting and crime from across Britain and the US, Welsh and Farrington (2008) find a statistically significant reduction in crime of 21% across street-lit areas compared with non-street lit controls. In reviewing the field research on street lighting, we cannot separate unconscious from conscious considerations, but by extrapolating from the laboratory evidence we can reasonably suggest that some component of the reduction in crime found when streetlights are present might be caused by proximate mechanisms interpreting light as a cue of visibility and causing behaviour change rather than reductions in crime in lit streets being caused only by explicit consideration of the increased chances of getting caught.

## 1. Introduction.

My preceding effort to put the watching eyes effect in an evolutionary context is the most extensive and comprehensive investigation into whether we should expect an eye cue and perceived surveillance effect and why this might be so that I am aware of. However, the idea that evolutionary psychology might explain the watching eyes effect is not new. The earliest studies of the watching eyes effect advance this idea, none better than Burnham and Hare (2007) who described the *evolutionary legacy hypothesis* to explain the watching eyes effect suggesting that behavioural anomalies in the presence of eye cues might be part caused by a mismatch between human ancestral conditions and the modern environment. They suggested, as I do ten years later, that it might be due to an adaptive proximate mechanism that makes us acutely sensitive to social gaze so much so that it could be triggered by a false cue – such as an image of an eye (Burnham, 2003; Burnham & Hare, 2007).

All things considered, the weight of evidence – ethological, evolutionary, cognitive, psychological, and neural – suggests a significant possibility that eye gaze might exert an influence on our behaviour. Taken in turn, the ethological evidence suggests there is an adaptive evolutionary advantage gained by those creatures that are able to detect the presence of predator or prey via the rapid identification of eye morphology. The evolutionary evidence suggests that human eyes evolved to communicate and that rapidly detecting the presence of others may have been adaptive in early human social groups to protect an individual's reputation and infer the intentions of others towards them. Lastly, cognitive and neuropsychological evidence suggests that our response to gaze is fast, at least partly automatic and unconscious, focuses significantly on emotional content, causes us to perceive objects differently, to process information differently, and that it elicits emotional arousal in the brain which exerts control over our actions via different neural pathways depending on whether or not we are observed.

All of which suggests that gaze should change behaviour. But the question remains: might we be so sensitive to gaze that eye cues alone can trigger such responses?

### 1.2.11. Eye-Images as Eye-Cues

The answer, it would appear, is that we might. Eye cues may function as a gestalt, a still partial image prompting the brain to consider the independent existence of the whole object via pattern recognition. In this case, the eyes cue consideration of the 'whole person' or a social presence (Kandel, 2012). Tsao, Freiwald, Tootell, and Livingstone (2006) and (Tsao, Moeller, & Freiwald, 2008) have shown how the brain processes face images along 6 main patches, about 3mm in diameter, arranged along an axis from the back of the inferior temporal lobe to the front. Examining each 'patch' under fMRI shows that they each respond to a different dimension or aspect of a face or face image, and activation of any one causes the others to activate, suggesting that we unconsciously create an overall picture of the face and entity rapidly even when we only have part of the face 'pattern' to work from (Freiwald & Tsao, 2010; Kandel, 2012; Tsao et al., 2006; Tsao et al., 2008).

The majority of the cells in the middle patches are specifically responsive to large irises (Freiwald & Tsao, 2010; Kandel, 2012; Tsao et al., 2006; Tsao et al., 2008), while electrophysiological studies have shown that eyes presented in isolation elicit a significantly greater neural (ERP) response than whole faces or any other feature of the face alone (Bentin, Allison, Puce, Perez, & McCarthy, 1996). This may explain why the eyes can be such a powerful cue: their mere presence cues us to construct a 'whole' social presence of another person (Freiwald & Tsao, 2010; Kandel, 2012; Tsao et al., 2006; Tsao et al., 2008), rapidly, and unconsciously, causing the changes in neural processing previously described.

Further evidence for our ability to identify and adapt to social presence comes from studies of human infants. Empirical evidence shows that infants are able to identify a three-spot pattern (two dots for the eyes, one for the nose or mouth – suggestive of a face) immediately upon birth and suggests that this pattern recognition is programmed and developed prenatally during REM sleep in the womb, thereby facilitating facial recognition and orientation to the eyes of adults from the moment of birth (Morton & Johnson, 1991). This proposition has received support from neuro-computational modelling that again suggests the response even to simple 3-dot eye cues is unconscious, automatic, innate, subcortical, and present from the moment we first open our own eyes to the world (Bednar & Miikkulainen, 2003).

That the response is unconscious is most powerfully suggested by the case of T.N. a physician who was left cortically blind following a series of strokes. Presented with a picture of a face with gaze averted and a second directly gazing out of the computer screen and back at him, he was blind to what he was looking at – unable to see the face or, once told it was in front of him, to guess which way it was looking. However, under fMRI, he showed significant coactivation of the amygdala and several brain areas related to face and gaze processing when presented with the direct gaze, but not the averted gaze image, remaining at all times consciously unaware of the stimulus. In other words, he showed a pronounced neural response to gaze cues without a conscious recognition that they had any influence on him at all (Burra et al., 2013).

Reviewing all of the evidence in this section we can see that our fixation on the eyes, our physical response to them, our cognitive ability to see things from the perspective of another's gaze, our fast, automatic, unconscious neural processing, our altered perceptions and motor-control mechanism, and our unconscious consideration of reward might all be activated by eye cues as they trigger a gestalt representation of the watching 'other' in the brain.

Much of psychology has, of course, long relied on photographs of eyes to generate physical, neural and cognitive reactions, accepting them as a substitute for human-to-human gaze cueing [for a review of the use of photos of eyes and faces in inducing and measuring *unconscious* responses in humans see V. Axelrod, Bar, and Rees (2015)]. The science of the 'watching eyes effect' then, examines whether these well-established unconscious, automatic, physical, cognitive and neural responses to the eyes of others have a behavioural manifestation when triggered by a false 'eye cue'.

### 1.2.12. Mere Presence: A Psychology of Being Watched

Also relevant to any research into the effect of surveillance on behaviour is that examining the effect not of gaze alone but of the presence of others, of audience effects, and it is here that the psychological study of surveillance effects has been most extensive.

In early studies of cycling, Triplett (1898) argued that the mere 'presence of another rider is a stimulus to the racer in arousing the competitive instinct; that another can thus be the means of releasing or freeing nervous energy for him that he cannot of himself release'. As such, his study of what he called dynamogenic factors in cycling might properly be accredited as the earliest empirical research into the effect of the presence of others – and perhaps of feeling watched – on human behaviour. However, psychological science exploring sensitivity to gaze did not begin in earnest until Zajonc's (1965) work on the effect of the 'mere presence' of others. The overriding (sorry!) finding that stemmed from this 'mere presence' research was that being observed appeared to enhance performance in well-learned responses but inhibit it in novel or poorly learned tasks (Geen, 1991).

The 'mere presence' hypothesis was further pursued in related work on 'evaluation apprehension' (Cottrell, 1968) where having others present inhibited rather than enhanced performance. Evaluation apprehension and mere presence came to be understood as *objective*, *public* self-awareness or *public* self-consciousness – to reflect that both represented changes in how one sees oneself when an audience is present (Fenigstein, Scheier, & Buss, 1975; Froming, Walker, & Lopyan, 1982; Govern & Marsch, 2001). A related, but distinct surveillance effect posited around the same time as the concept of objective self-awareness was first being discussed was that of '*private* self-awareness' - where observation prompts us not so much to worry about what others think of us, but rather prompts us to consider whether we are acting in alignment with the way we see ourselves (Buss, 1980; Fenigstein et al., 1975; Froming et al., 1982; Wicklund & Duval, 1971, 1972).

To demonstrate the difference, Froming et al. (1982) used a mirror to induce private self-awareness and an audience to induce public self-awareness in participant ''teachers''. The "teachers" posed problems to the co-participant "learner" who was separated from them by a wooden partition. If the co-participant gave an incorrect response the "teacher" must administer a punishment to the learner by selecting an electric shock from 1-10. All participant teachers were pre-selected having indicated on a questionnaire some months prior to the study that they believed physical punishment could sometimes be justified. However, on the day all were instructed only to try to teach the answers to the problems posed to the learner as effectively as possible. Participant "teachers" who were in the audience condition had a further two co-participants sat either side of them on their side of the wooden partition and were told that they would be evaluating their performance as a teacher - to induce audience effects, public self-consciousness and evaluation apprehension. Participant "teachers" in the private self-consciousness condition had a mirror placed at eye level in front of them throughout the experiment. Participants who were primed to be privately self-conscious shocked at significantly lower levels than those with an audience. Froming et al. suggest this is because those primed to be privately self-conscious were living up to their view of themselves as 'good people', whereas those primed to be publicly self-aware were living up to standards their evaluators would approve of.

## 1. Introduction.

Public self-awareness and private self-awareness have become key concepts in the field of social cognition – which seeks to understand how the processes of thinking about or in the presence of others differs from the processes underlying our thinking about or in the presence of objects (Fiske & Taylor, 2017, pp. 20-21, 28). Both can be activated by being watched by others prompting self-stereotyping – it's just the point of reference that changes (Abrams, 1999, pp. 219, 220-221). Private self-conscious is also activated by the presence of a mirror, a photo, or a video of oneself, and, as described above, promotes behaviour consistent with internal attitudes. In contrast *public* self-consciousness promotes a focus on external aspects of the self, that is, how one appears to others, and encourages behaviour aligned with group norms and expectations (Hass, 1984; Wiekens & Stapel, 2010). In light, then, of the general consensus among personality theorists that people are a collection of selves, or schemas activated by cues (Deci & Ryan, 2002, p. 4; Mischel & Shoda, 1995) *often automatically and unconsciously* (Fiske & Taylor, 2017, pp. 34-45) eye cues might 'work' by prompting us to 'see' ourselves, activating one of two schemas adaptively suited to thrive in the spotlight of social gaze: either cueing us to live up to our expectations of ourselves, as advocates for *private* self-consciousness argued, or prompting us to try to conform to social norms and meet the expectations of others, as Foucault and advocates for *public* self-consciousness might hold, or perhaps a combination of both.

At this point it should be noted that Zajonc's mere presence hypothesis, though still prevalent, has received inconsistent empirical support. While some studies have found that the presence of an observer changes behaviour (e.g. H. Markus, 1978; Schmitt, Gilovich, Goore, & Joseph, 1986), sometimes decreases performance levels (Forgas, Brennan, Howe, Kane, & Sweet, 1980), and often increases generosity and prosociality (e.g. Fujii, Takagishi, Koizumi, & Okada, 2015; Kurzban, 2001), others have found that the presence of an observer has limited effects (e.g. Abrams & Manstead, 1981; Cacioppo & Petty, 1986) or is only effective when the observer is from the same 'group' as the observed (an '*in-group*', in the jargon) (Abrams, Wetherell, Cochrane, Hogg, & Turner, 1990). Nevertheless a recent meta-analysis has found that the presence of others does indeed correlate with a small increase in prosocial behaviour (Bradley, Lawrence, & Ferguson, 2018) and recent studies have found that the mere presence of an observer also correlates with an eclectic array of behavioural change, including: inhibiting task performance through evaluation anxiety (Yoshie, Nagai, Critchley, & Harrison, 2016); reducing contagious yawning (A. Gallup, Church, Miller, Risko, & Kingstone, 2016); magnifying people's perception of their performance – both errors and successes - and their assessment of their contribution towards team successes and failures (Steinmetz, Xu, Fishbach, & Zhang, 2016); increasing risk aversion in gambling (Lemoine & Roland-Lévy, 2017); improving problem solving performance (Laird, Bailey, & Hester, 2018); and (as previously described) influencing primate behaviour – improving performance in monkeys in a visuomotor task (Demolliens et al., 2017).

### 1.2.13. Research Overview

There still exists a sizeable gap in the literature. Although the evidence suggests we might expect an effect of gaze and surveillance on behaviour, and that any effect might be triggered by false cues, it provides little understanding of the boundary conditions of such an effect. It can give little guidance on when, and under what circumstances, the effect might or might not be found.

Consequently, the following chapters explore the watching eyes effect literature in more detail and document a series of our own studies into the phenomenon. We include in these studies the addition of the 'mere presence' of a camera to understand whether any effect is dependent on the morphology of the human eye or whether it might translate to the kind of ubiquitous surveillance that increasingly characterises societies globally.

## 1. Introduction.

Chapter 2 reports the first study, examining the effect of watching eyes or the mere presence of a camera – henceforth referred to collectively as ‘surveillance cue effects’ - on prosocial behaviour.  Chapter 3 follows this with a meta-analysis and review of the watching eyes effect on antisocial behaviour.  In Chapter 4 surveillance cue effects on antisocial behaviour are explored in three large-scale studies: the first under laboratory conditions and the second two online.  In Chapter 5 we report the results of a field experiment exploring the watching eyes effect on a range of criminal and antisocial behaviours in the British town of Hereford.  Chapter 6 reports a carefully designed replication of what has been perhaps the most influential watching eyes experiment on public policy - Nettle et al.’s (2012) study showing how watching eyes may reduce cycle crime (Nettle et al.’s study is described in more detail in Chapters 3, 5 and 6) – while Chapter 7 draws conclusions from the studies and evidence reported herein, and outlines its implications for public policy and future research.

## 1. Introduction.

# Chapter 2

# 2

# Study 1 Perceived Surveillance & Prosociality

> *"The eyes of the Lord are everywhere, keeping watch on the wicked and the good."*
>
> Proverbs 15:3

Chapter 2 presents Study 1 in which we explore the effect of the presence of eye cues or a webcam or prosocial behaviour in two tasks. The first task tests whether participants help others more frequently when eye cues or a camera are present through a simple computer game. The second uses the dictator game paradigm to test whether participants donate more to others when eye cues or a camera are present. Additionally, we explore whether participants report higher levels of public self-awareness in the presence of eye cues or camera as a proxy for how much they felt watched. We find no effect of eye cues or a camera on helping behaviours in the computer based task, nor on donations in the dictator game, nor on levels of public self-awareness.

# 2. Study 1 Perceived Surveillance & Prosociality

## 2.1. Introduction

Chapter 1 showed that there is reason to think humans may be highly sensitive to cues of being watched - so much so that false cues might trigger behavioural responses automatically and unconsciously.

This chapter describes an experimental investigation into the effect of perceived surveillance on prosocial behaviour. In this study three dependent variables were selected, (1) how frequently participants helped others in a computer-based task, (2) how much they donated to others in an economic game, and (3) how 'watched' participants felt. The study tested whether these dependent measures varied by surveillance condition, the independent variable, which had four levels (1) a no surveillance control (2) the presence of 'watching eyes' displayed in a poster, (3) the mere presence of a camera, and (4) the highly salient presence of a camera. A main effect of surveillance cues on prosocial behaviour was expected, with participants in the eyes, camera and camera salient conditions being expected to help more and give more than those in the no surveillance cue control.

One problem with much of the 'watching eyes' literature is the absence of ecological validity in many experiments. The findings of economic or social preference games such as the Dictator Game, Ultimatum Game and Trust Game do a poor job of explaining or predicting social behaviours in field experiments or people's self-reported social behaviours in everyday life (Galizzi & Navarro-Martinez, 2018; Levitt & List, 2007a, 2007b; Voors, Turley, Kontoleon, Bulte, & List, 2012).

For this reason, we selected the Zurich Prosocial Game (ZPG) as used by Leiberg, Klimecki, and Singer (2011) to try and increase the ecological validity of economic games.

Leiberg et al. employed this task to measure the effectiveness of various compassion and memory interventions in promoting prosocial behaviour, as indicated by how often participants helped another participant in their computer-based task when there was no incentive do so.

In the task participants are explicitly told that their goal is to navigate their on-screen icon through a maze to a treasure chest and will be paid if they make it within a given time limit. Although they can help the other on-screen player, who is navigating along a separate path to a separate treasure concurrently, there is no reward for doing so. In Lieberg et al.'s study (2011), participants had undergone differing forms of compassion and memory training. The authors found significant differences in how often participants helped between those that had undergone compassion and memory training and those who had not. Differing forms of compassion and memory training also correlated with differing helping rates – suggesting that some forms of compassion and memory training were more effective in promoting prosociality than others. They judged the task to be a sensitive and reliable measure of prosocial behaviours.

We decided therefore that this task would be appropriate to assess the effectiveness of surveillance interventions in promoting prosocial behaviour. In contrast to field experiments, it allowed us to control the environment to better isolate effects. In economic games the only thing participants have to do is decide to give or take money, which may lead to demand effects. To overcome this, the ZPG makes helping – the principal outcome measure - incidental to the goal of the task. By divorcing the dependent variable (helping behaviour) from the participant's goal in the task (making money) it more closely resembles everyday life in which helping others is usually something done concurrently with pursuing one's own goals.

In addition to the Zurich Prosocial Game we employed the Dictator Game, the most widely used measure of generosity in watching eyes experiments (see Nettle et al., 2013 & Northover et al. 2017 for reviews).

When designing our surveillance conditions, we aimed to create as ecologically valid an experimental design as is possible under laboratory conditions. In public life we do not know if the cameras we pass, on the street, in stores, or outside private residences etc, are recording or not. Van Bommel (2012, 2014) presents some evidence that the mere presence of a camera can influence behaviour. He demonstrated that the presence of a webcam correlated with participants being more concerned with managing the impression they made on others, and reported significantly higher self-monitoring scores with a camera present than in its absence. Similarly, switching on a webcam induced participants to assist more in an online help-seeking forum where there was no incentive to do so. From these results, Van Bommel et al. (2012) conclude that the subjective *feeling* of public self-awareness may be enough to increase helping behaviours.

In a follow-up study in 2014 Van Bommel et al. replicated the well-known bystander effect: the more people that are present at the scene of an accident or crisis the less likely each individual is to intervene or help (Fischer et al., 2011; Latané & Nida, 1981). In Van Bommel et al.'s study, participants witnessed the theft of money from the experimenter by what they believed was another participant (but in reality, was a confederate of the experimenter). Having replicated the bystander effect, showing that the more people were present the less people helped, they demonstrated that by placing a CCTV camera in the room they could reverse this effect – when the camera was there people were *more* likely to help when others were present than when they were alone, despite the fact that fewer than half (45%) of participants were able to recall consciously seeing the camera. As with their first study, Van Bommel et al. again found that participants reported significantly higher self-awareness scores in the presence of a camera. These studies provide tentative evidence for the theory behind Bentham's panopticon concept – that man need not necessarily be watched, but just 'conceive himself to be so' in order to behave more prosocially (see Chapter 1). In the current study, we aimed to examine whether the mere presence of a camera compared to conspicuous and explicit filming made a difference to the degree of prosocial behaviour observed.

Some previous research has looked at the effect of cameras on prosocial behaviour. Buchanan, McMahon, Simpson, and Wilson (2017), for instance, found no effect of camera presence on Dictator Game decisions – even when participants knew that recorded footage of their decisions was being filmed and screened live on a monitor to a group of observers. However, in Buchanan et al.'s study the recipient of the dictator's 'donation' was visible to the dictator at the time she/he decided how much, if anything, to give to them. Therefore, participants in the control condition were not anonymous and dictators were under a form of surveillance at the time they made their decision, (albeit there was 40'/12.2m distance between them and the recipient). This mutual visibility meant that there was no 'no surveillance' condition to act as a control, which may have been an important confound – participants were watched in all conditions. In support of this interpretation, Buchanan et al. (2017) note that the magnitude of donations was higher in this study than is typical in dictator game studies that provide anonymity. It is possible that all donations were elevated by the direct surveillance of the dictators so that cameras had no additional effect. In our study we aim to see if just the presence of the camera has the same effect in the absence of other people.

The present study also explores a range of individual differences measures based on those identified in Chapter 1 as potentially influencing responses to surveillance cues. We were particularly interested in exploring self-awareness as a measure, anticipating a main effect of Surveillance Condition on self-awareness: such that those that most actively self-monitored and adjusted their behaviour according to the situation would likely be most affected by surveillance cues. It was only later that we became aware that Pfattheicher & Keller's work on public self-consciousness and eye cues had indeed found evidence for this possibility (2011; 2015). We discuss this work here as it usefully informs this study. We then seek to replicate their findings using the same measures in Chapter 4.

### 2.1.1. Public Self-Awareness, Public Self-Consciousness & Eye Cues

Given the extensive literature examining the watching eyes effect and the controversy that surrounds it, it is surprising that little effort has been made to establish that the presence of watching eyes makes people feel more watched. Rather most research has proceeded from finding a correlation between the presence of an eye cue and a change in some behaviour - such as increased generosity or reduced cycle theft – to the assumption that eye cues made people feel watched. A further extension assumes that this feeling of being watched generates unconscious reputational concerns, and that these concerns are what cause changes in behaviour (e.g. See the 'evolutionary legacy hypothesis' in Burnham & Hare's foundational watching eyes article (2007)). While this is logical and reasonable, the evidence for it is indirect and is therefore open to challenge.

Pfattheicher and Keller (2015) sought to fill this lacuna. They demonstrated that people over-estimated the extent to which others had been looking at them when asked to record how much they had been watched by others in the presence of watching eyes compared to a 'no eyes' control. The increased 'spotlight effect' (Gilovich, Medvec, & Savitsky, 2000) was taken to demonstrate that people did feel more watched in the presence of eye cues.

Pfattheicher and Keller (2015) extended this research in a second experiment finding that individuals with high public self-consciousness (Fenigstein et al., 1975) - that is those especially concerned with their appearance and reputation in the eyes of others - were more strongly affected by the presence of eye cues. They conclude that watching eyes do indeed make people feel more watched.

Although at this stage of our research we were not aware of Pfattheicher and Keller's (2015) studies we did anticipate that measures of the extent to which a person tries to actively manage the way they present themselves in public may moderate any effect of eyes, or alternatively that the presence of eye cues or a camera might lead people to report a greater sense of the need to manage the way they present themselves. To this end, we employed the Snyder-Self Monitoring Scale (Snyder, 1974), which examines precisely this - the extent to which a person actively manages the way they present themselves in public.

### 2.1.2. Personality Measure – HEXACO PI-R

Given the variety of personality dimensions said to influence responsiveness to eye cues we added the HEXACO PI-R (Ashton & Lee, 2009) to look for interaction effects between surveillance condition and personality on donations and helping behaviour. The HEXACO PI-R is the most reliable global measure of personality (Ashton & Lee, 2018) and measures participants on six personality dimensions: Honesty-Humility, Emotionality, eXtraversion, Agreeableness (vice Anger), Conscientiousness and Openness to Experience (HEXACO). In addition to its comprehensiveness, our selection of the HEXACO PIR was also influenced by the findings of Hilbig, Thielmann, Hepp, Klein, and Zettler (2015) who noted that those high in Honesty-Humility (HH) consistently gave more in Dictator Games.

We expected a main effect of HH on generosity with higher HH scores relating to more generosity. We also expected to find an interaction between HH scores and surveillance cues. Low HH participants should be the least generous in the control, but the most responsive to surveillance cues. Their inclination to help would be instrumental, motivated by a desire to enhance their reputation to gain future benefit. Consequently, they would help less in the control than in the eyes and camera conditions. There would be no interaction by condition for those with higher HH scores since their desire to help would be a function of their intrinsic pre-disposition rather than motivated by concerns regarding reputation management and future benefit.

## 2.2. Method

### 2.2.1. Study Sites & Participants

Testing was conducted across 6 military sites and at one Oxford-based civilian commercial company between May and November 2016. Excepting those at the civilian company, most were serving military personnel, although at all sites a number of participants were non-military civilians and civil servants employed on the base. Our sample size was 207, Control *N* = 49, Eyes *N* = 57, Camera *N* = 53, Camera Salient *N* = 49.

### 2.2.2. Measures

We measured the frequency with which participants helped the other player in the ZPG and how much participants donated in the dictator game as our primary outcome measures. Following completion of the ZPG and dictator game, all participants completed two questionnaires, the Snyder Self-Monitoring Scale to look for a main effect of surveillance condition on how much people felt watched and the HEXACO-PI-R to look for interaction effects between personality dimensions and condition on our dependent variables. Participants also provided their ethnicity and were asked to say whether they were religious, not religious, or preferred not to answer.

The ZPG allows custom options. Here we describe the options we selected, and the ZPG task in detail, to enable replication. In our version of the task participants navigate a virtual character through a maze (see Figure 2.1.) to try to reach a treasure chest in a limited amount of time. Each treasure is worth £0.50. At the same time, participants see the virtual character of an ostensible co-player who they believe is situated either in another room at the testing site or at some other site within Oxford University where, they are told in the pre-briefing material, 'the experiment is being conducted concurrently'. This co-player is also trying to reach their own treasure along their separate path. Note that the two players do not share the same paths in the maze and do not compete for the same treasure. Thus, in principle, the game can be played while completely ignoring the other player. Participants are told that in each round of the game they are connected via the internet with a new co-player different to those they have been paired with previously and those they will see in future rounds. In reality, the co-player is always controlled by the computer in order to ensure participants are responding to consistent stimuli.

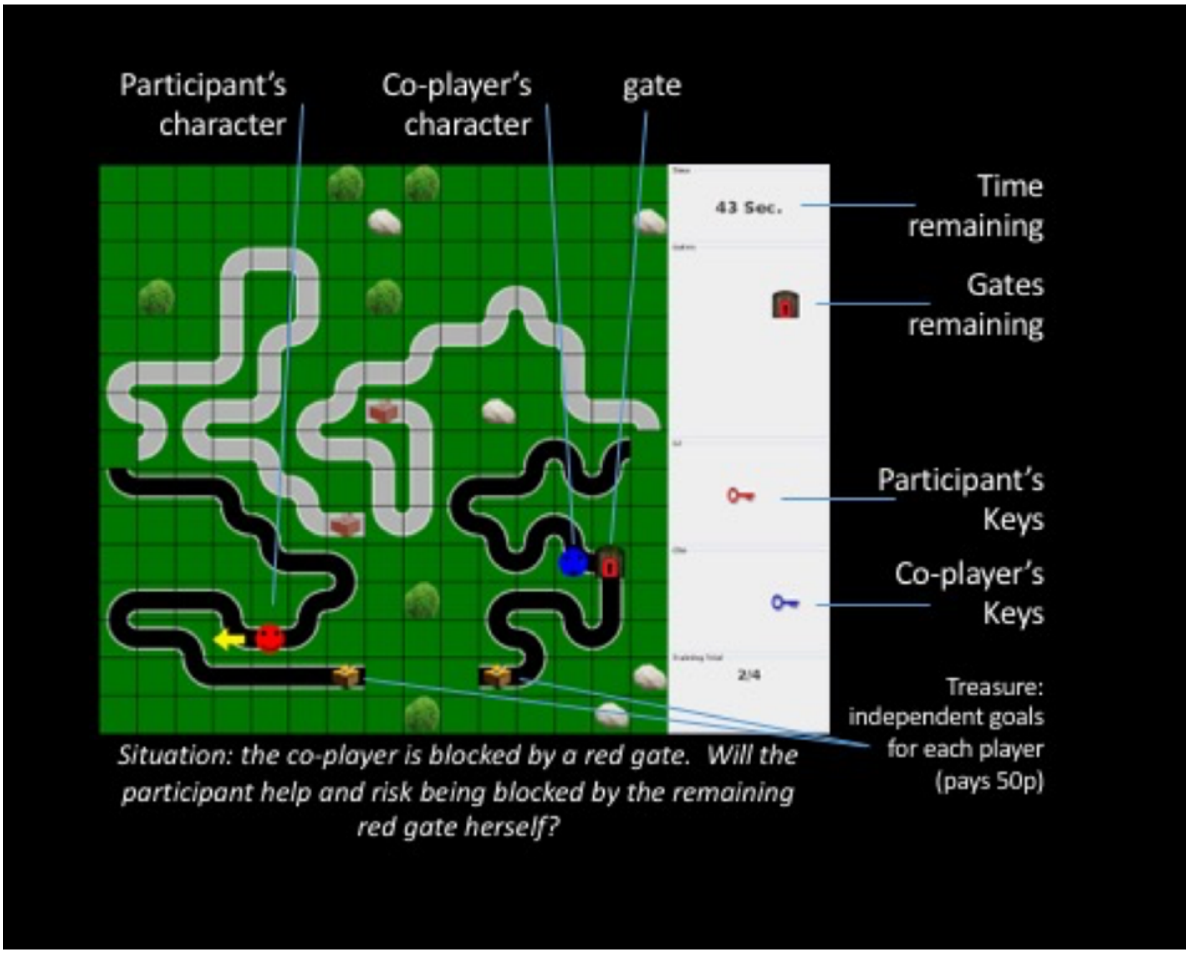

Participant's character
Co-player's character
gate
43 Sec.
Time remaining
Gates remaining
Participant's Keys
Co-player's Keys
2/4
Treasure: independent goals for each player (pays 50p)
Situation: the co-player is blocked by a red gate. Will the participant help and risk being blocked by the remaining red gate herself?

*Figure 2.1. Labelled screenshot from the Zurich Prosocial Game (ZPG).*

At the onset of each round four paths to the treasure are displayed, two for each player. The participant and the co-player select one of the two paths open to them. The players move their virtual character along their chosen path. Red and blue gates appear on all four paths. If they fall on the paths that aren't selected by either player, they are irrelevant. However, if they fall on the selected paths they *are* relevant and when this happens they can block the participant or the co-player from proceeding further. Each of the two players is equipped with red and blue keys with which they can open the correspondingly coloured gates. When the co-player runs out of keys, participants can use their own keys to open the gates for them. Importantly, participants cannot delay their help to observe the progression of the game (i.e. whether they need their keys themselves) because a blocked character becomes permanently inactive after a few seconds being stuck behind the gate. Consequently, a participant has to unlock a gate - her own or that of the other player - within this time limit. During each trial, participants can see how many gates are still going to fall but they do not know which paths they will fall on. They also know the colour and the number of keys that they and their co-player have, and how much time is left (again, for a screenshot of the game display, see Figure 2.1.).

When playing the game, participants wear headphones and sounds add emphasis to events on the screen (e.g. when a gate is falling). We disabled the ‘distress cue’ feature of the game used by Leiberg, Klimecki & Singer, where a sound of crying would sometimes play when the co-player became trapped behind a gate, as this may have confounded the results perhaps eliciting greater empathy in some trials than others. Importantly, to reduce demand effects, participants are not told that the purpose of the game is to help the co-player. Instead, they are instructed that the aim of the task is to reach a goal - the treasure - to optimize monetary winnings.

Participants completed 11 experimental trials drawn at random from a combination of: no-reciprocity or reciprocity trials, with high-cost and low-cost and fair or unfair trials. We explain each of these variations in turn and then list how they were combined to create 11 trials.

In no-reciprocity trials participants could help their co-player if they wanted but could see that there would be no opportunity for their co-player to help them in the future. This could be because there were no further gates to fall, because their co-player did not have the appropriate key to open the gates that were to fall, or because their co-player simply had no more keys.

In reciprocity trials participants had the opportunity to help their co-player after their co-player had earlier helped them by opening a gate behind which they had been stuck without a key. In both reciprocity and no-reciprocity trials, participants could also see that there would not be any further opportunity for the co-player to reciprocate if they themselves chose to help. By designing the trials this way and by ‘changing’ the co-player for each trial we excluded the possibility of participants helping because they anticipated that they might need the co-player to reciprocate later on.

Reciprocity or no reciprocity trials could be either low- or high-risk. In the high-risk variant, participants could see which gates were still to fall and which keys they had left, and so knew that after they helped the co-player, there would be a 25% chance that they would need the donated key to reach the treasure themselves; in the low-risk variant, players could see there were no more gates to fall that they would be unable to open and so knew that they could donate keys without risk of needing them later themselves. In low-risk trials the only cost to participants from helping was therefore a loss of time.

Trial types could be fair or unfair. There were fair if the co-player first helped the participant. They were unfair if the co-player did not help the participant, in which case the participant’s virtual character ‘dies’ but they still have the chance to help the co-player by unlocking a gate for them later.

Trial types appeared in random order with the restriction that the first reciprocity trial could appear at the end of trial three. This restriction was introduced to reduce the likelihood of anchoring effects that might have followed if a participant encountered a helpful co-player right away and thus saw helping as the ‘norm’.

The 11 trial types presented in random order with the constraint that the first reciprocity trial must not appear before trial 3 were comprised of:

- 1 x No Help trial. A trial in which the two players do not need help.
- 2 x Low Risk, Non-Reciprocity trials. Trials in which the participant may help the ostensible co-player without risk of failing to reach the treasure, knowing the co-player cannot help them later.
- 2 x High Risk, Non-Reciprocity trials. Trials in which the participant may help the ostensible co-player, using a key the participant might need later to reach the treasure, knowing the co-player cannot help them later.
- 2 x Reciprocity, Low Risk, Fair trials. Trials in which the ostensible co-player helps the participant before needing help in return. The participant may help the ostensible co-player without risk of failing to reach the treasure.
- 2 x Reciprocity, High Risk, Fair trials. Trials in which the ostensible co-player helps the participant before needing help in return. To help the participant must use a key they might need to reach the treasure themselves.
- 2 x Reciprocity, Unfair trials. Trials in which the ostensible co-player refuses to help the participant before needing help from the participant to continue. The participant's virtual character dies but still has the chance to help the ostensible co-player.

At the beginning of the game, participants were first given written and verbal instructions and asked five questions to test their comprehension of the task. Participants then completed four practice trials to familiarise themselves with the run of the game and to determine individual reaction time thresholds. To offset individual differences in speed and proficiency with computer games, the individual time limit for all trials of a given game was set at the average time the individual required to reach the treasure in their practice trials plus 5 seconds.

In our Dictator Game participants found on the desk £5.00 comprised of 2 x 50p, and 4 x £1.00 coins. Additionally, they found two envelopes one marked ‘for you to take-away’, the other marked ‘for a random participant’. Participants were told in advance that their task was to allocate the money to each envelope as they wished, placing the envelope marked ‘for a random participant’ in a plain black post-box located on the desk and then taking the envelope marked ‘for you to take-away’ with them. The envelope marked ‘for a random participant’ was also marked with the participant’s randomly allocated, double-blind, participant number to enable the cross-referencing of donations by condition upon completion of the experiment.

It was made clear to participants that the experimenters had no way of linking participant numbers back to specific individuals - and therefore their donations or how much they kept were completely anonymous. A suspicion check was undertaken upon completion of the experiment to check that participants had believed this pre-brief. Participants were asked whether they believed their decisions were anonymous at all times throughout the study.

### 2.2.3. Study Design

Participants were randomly allocated to one of four conditions: A no surveillance cue control, Eye Cues, Camera (Present), and Camera (Salient). We selected the ‘watching eyes’ and control posters used by Nettle, Harper, Kidson, Stone, Penton-Voak and Bateson (2013) as the eye cue in our experiment since they had found that the presence of these cues led to increased probability of donating something rather than nothing (80% donated in the ‘eyes’ condition compared to 63% in the control). We modified these by replacing the Newcastle University logos with those of Oxford. For the control posters we replaced The Institute of Neuroscience Logo with the crest of Magdalen College, Oxford (see Figure 2.2.).

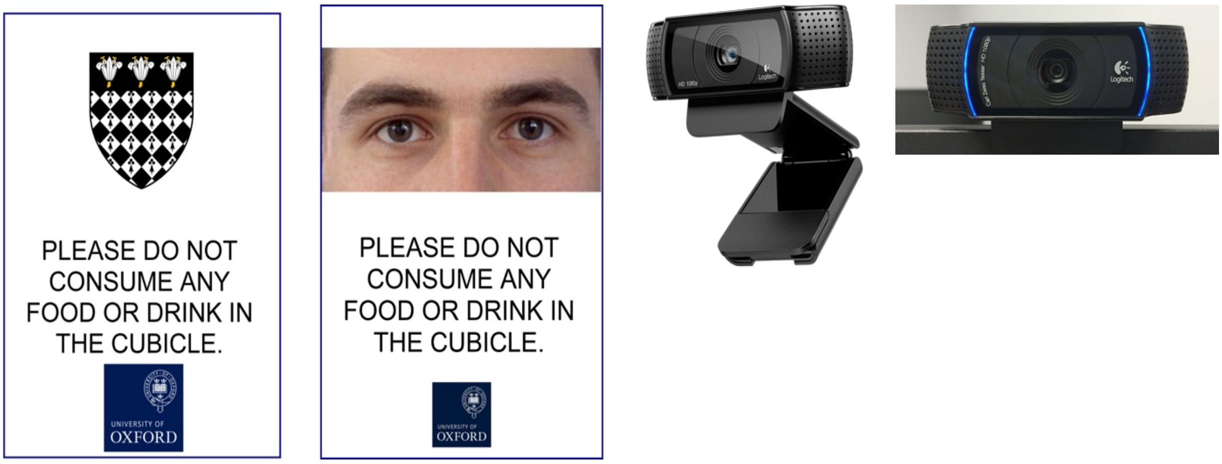

PLEASE DO NOT CONSUME ANY FOOD OR DRINK IN THE CUBICLE.
UNIVERSITY OF OXFORD
PLEASE DO NOT CONSUME ANY FOOD OR DRINK IN THE CUBICLE.
UNIVERSITY OF OXFORD
HD 1080p
Logitech
Logitech

*Figure 2.2. Conditions (from left) Control, Eye Cues, Camera (Present), Camera (Salient).*

For the camera conditions we selected the Logitech C920 HD Pro USB 1080P Webcam. In the Camera (Present) condition the camera was set up and pointing at the participant, but the participant was not told it was there, it was not switched on, and displayed no lights. Since participants could not know for sure it was not recording this simulated the way we usually interact with surveillance cameras in everyday life. In the Camera (Salient) condition the camera was in the same location, but switched on, which meant that a blue light illuminated on the camera to draw attention to its presence. Participants in the camera salient condition were told that a camera would be present, but it was explained that it would record only their faces and upper torso, thus their responses on the computer and their allocation decision in the Dictator Game would remain anonymous. To ensure that participants consciously believed they were anonymous we conducted a suspicion check on completing the study, asking participants in all conditions whether they had believed that their decisions were anonymous at all times.

### 2.2.4. Procedure

Testing locations varied in size but were uniformly plain-walled, white-washed rooms with strip-lighting, a single desk, chair and testing computer. Any windows were temporarily papered over.

Research assistants varied by location. Sometimes junior officers or non-commissioned ranks assisted with the study, at other times research assistants were recruited from within the Department of Experimental Psychology at the University of Oxford. Military personnel employed as assistants were never in uniform to minimise any influence that rank or collective identity might exert. Research assistants were asked to declare if they knew participants in advance of their arrival to ensure participants were never escorted in and out of the experiment by a friend or colleague in order to avoid any demand effects this might have. Upon arrival participants were given a written brief to explain the task and then had the opportunity to ask the research assistant any questions they had regarding the research. On some occasions due to the high number of participants volunteering, the lead researcher would provide participants with a pre-printed written brief and take questions on the study in private briefing room, before the assistant took the participant to the testing room.

Following the briefing participants were asked to sign a consent form if they wished to participate. Participants were then allocated a participant number, which was selected at random from a random numbers table by the research assistant. They were then taken to the testing room. After testing they were debriefed and paid by the lead researcher, who received only their treasure ‘score’. Since the amount of treasure attained was not the variable of interest and we could not know from this data how much they had helped or not, this procedure did not compromise participants’ anonymity.

Response data in the ZPG was collected digitally during participants completion of the computer-based task. Dictator Game allocations to ‘a random future participant’ were recorded at the end of each day by the lead researcher after unlocking the ‘post-boxes’ in the testing rooms. Each participant’s randomly generated participant number was written unobtrusively on the envelope containing their allocation for a ‘random future participant’, but participants were aware that the lead researcher could not link this back to any specific individual, while the research assistants would never see the envelope. Thus the anonymity of their decisions and allocations was preserved.

### 2.2.5. Data Plan

Preliminary analysis - the calculation of individual difference scores, mean donations and mean helping, as well as removal of incomplete or corrupted responses – was conducted in Excel. All further data exploration, transformation and analysis was undertaken in *R* (R Development Core Team, 2016). We tested the main effect of condition on mean decisions to help in the ZPG and mean donations in the Dictator Game by means of a series of-one-way ANOVAs. We also examined interaction effects with personality dimensions in the HEXACO-PIR, specifically with Honesty-Humility and with self-monitoring scores in the Snyder Self-Monitoring Scale.

All data and code are available at Open Science - see note at Appendix D at the end of this thesis for details. Appendix A at the end of the thesis contains a brief discussion and analysis of religiosity and ethnicity variables on which data was collected but which are not reported in the main write-up due to inconsistency in participants’ free-text responses.

## 2.3. Results

Our sample was less ethnically diverse than we would have wanted. Over 90% of our participants were white European (88% white British, 0.5% Irish, 2% ‘white other’). The remainder were 3.5% Black/African/Caribbean/Black British, 1% Indian, 1% other. Four percent did not respond. Participants were aged between 18 to 62, *M*=35.37, *SD*=10.12, *CI*=[34.00, 36.73]. Gender was not a variable of interest and was not recorded. Table 2.1. below summarises the main effects of surveillance condition on our dependent variables (helping in the ZPG, donations in the Dictator Game, Self-Monitoring Scores) and potential moderators (HEXACO PI-R personality dimensions) each in a one-way ANOVA. All of our participants confirmed (correctly) that they were confident their decisions remained anonymous. Preliminary analyses confirmed that there were no violations of the assumptions of normality, linearity, multi-collinearity, and homoscedasticity.

As Table 2.1. shows, there was no significant influence of condition on rates of helping in the ZPG, $F$ (3, 203) = 0.20, $p$ = 0.90, $\omega^2$ = 0.01 or generosity in the Dictator Game, $F$(3, 203)= 0.55, $p$ = 0.65, $\omega^2$ = -0.14.

Table 2.1. and Figure 2.3. show a number of main effects of condition on HEXACO PI-R personality dimensions. Participants reported lower levels of emotionality in the control condition (*M* = 2.64, *SD* =0.59) than when surveillance cues were present, (Eyes *M* =3.05, *SD* = 0.67; Camera *M* = 3.05, *SD* = 0.63; Camera Salient *M* = 2.91, *SD* = 0.53), a one way ANOVA showed that the differences were significant, *F(3, 203) = 3.18, p = .03,* $\omega^2$ = 0.03. Figure 2.3 shows mean emotionality scores by condition. Welsh two sample t-tests showed that emotionality was significantly lower in the control condition than in Eyes, $t$(97.75) = -3.20, $p$ = 0.002, $r$ = .32 and Camera Salient $t$(89.82) = -2.49, $p$ = 0.01, $r$ = .24 conditions, but not significantly lower in the Control than in the Camera Condition, $t(87.19) = -1.56$, $p = 0.12$, $r = .16$.

**Table 2.1.**

**Summary Statistics for Study 1: One-Way ANOVAs with surveillance condition as the independent variable, incidence of helping as the dependent variable and psychometric measures as covariates. Values given represent the group mean and standard deviation, *F* and *p* values are associated with the significance test for those means**

| | *Surveillance Condition* | | | | | |
|---|---|---|---|---|---|---|
| | *Control* | *Eyes* | *Camera* | *Camera Salient* | *F* | *p* |
| *N* | 49 | 57 | 53 | 48 | | |
| Helped – ZPG (mean (*sd)*) | 0.50 (0.24) | 0.55 (0.23) | 0.53 (0.22) | 0.56 (0.24) | 0.55 | 0.65 |
| Donated – Dictator game (mean (*sd*)) | 5.16 (2.94) | 4.95 (3.30) | 4.72 (2.88) | 5.04 (3.02) | 0.20 | 0.90 |
| Snyder Self-Monitoring (SM) Score (mean (*sd*)) | 11.65 (3.97) | 11.61 (3.59) | 10.64 (4.27) | 10.73 (4.30) | 0.96 | 0.41 |
| Honest-Humility (mean (*sd*)) | 3.72 (0.64) | 3.45 (0.64) | 3.57 (0.64) | 3.65 (0.66) | 1.70 | 0.17 |
| Emotionality (mean (*sd*)) | 2.64 (0.59) | 3.05 (0.63) | 2.87 (0.89) | 2.91 (0.53) | 3.18 | **0.03*** |
| Extraversion (mean (*sd*)) | 3.65 (0.52) | 3.40 (0.67) | 3.65 (0.59) | 3.56 (0.51) | 2.24 | 0.09 |
| Agreeableness (mean (*sd*)) | 3.44 (0.61) | 3.09 (0.57) | 3.08 (0.63) | 3.15 (0.56) | 4.06 | **0.01*** |
| Conscientiousness (mean (*sd*)) | 3.80 (0.56) | 3.63 (0.53) | 3.85 (0.50) | 3.94 (0.55) | 3.17 | **0.03*** |
| Openness (mean (*sd*)) | 3.55 (0.64) | 3.17 (0.82) | 3.38 (0.75) | 3.45 (0.66) | 2.58 | 0.06 |

* = significant at alpha < .05, ** = significance < .01, *** = significance < .001

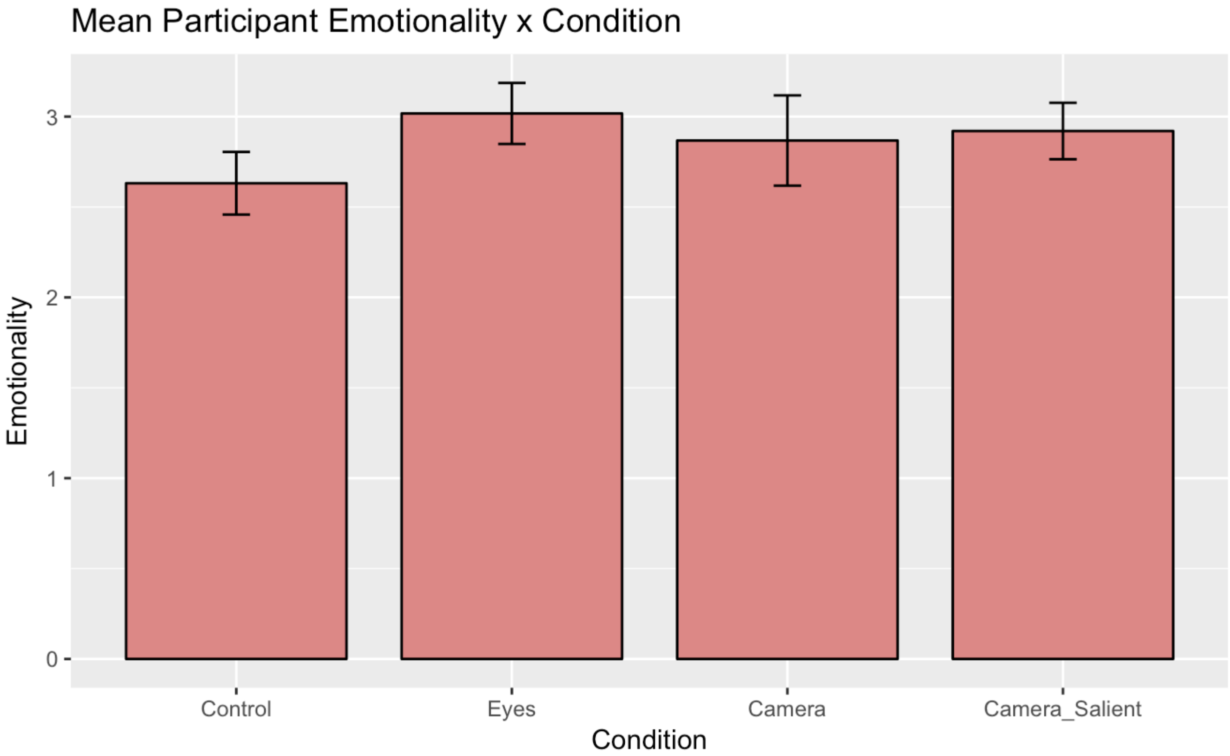


*Figure 2.3. Study 1. Mean HEXACO PI-R Emotionality Scores by Condition. Error bars represent the 95% Confidence Interval.*

Participants reported higher levels of agreeableness in the control condition (*M* = 3.44, *SD* =0.61) than when surveillance cues were present, (Eyes *M* =3.09, *SD* = 0.57; Camera *M* = 3.08, *SD* = 0.63; Camera Salient *M* = 3.15, *SD* = 0.56). A one way ANOVA showed that the differences were significant, *F(3, 203) = 4.06, p = .008,* $\omega^2 = 0.04$.

Figure 2.4. shows mean agreeableness scores by condition. Welsh two sample t-tests showed that agreeableness was significantly higher in the Control condition than in each of the surveillance cue conditions respectively, Eyes, *t*(99.10) = 2.30, *p* = 0.003, *r* = .29 ; Camera *t*(99.72) = 2.90, *p* = 0.005, *r* = .28; Camera Salient *t*(94.41) = 2.47, *p* = 0.02, *r* = .25.

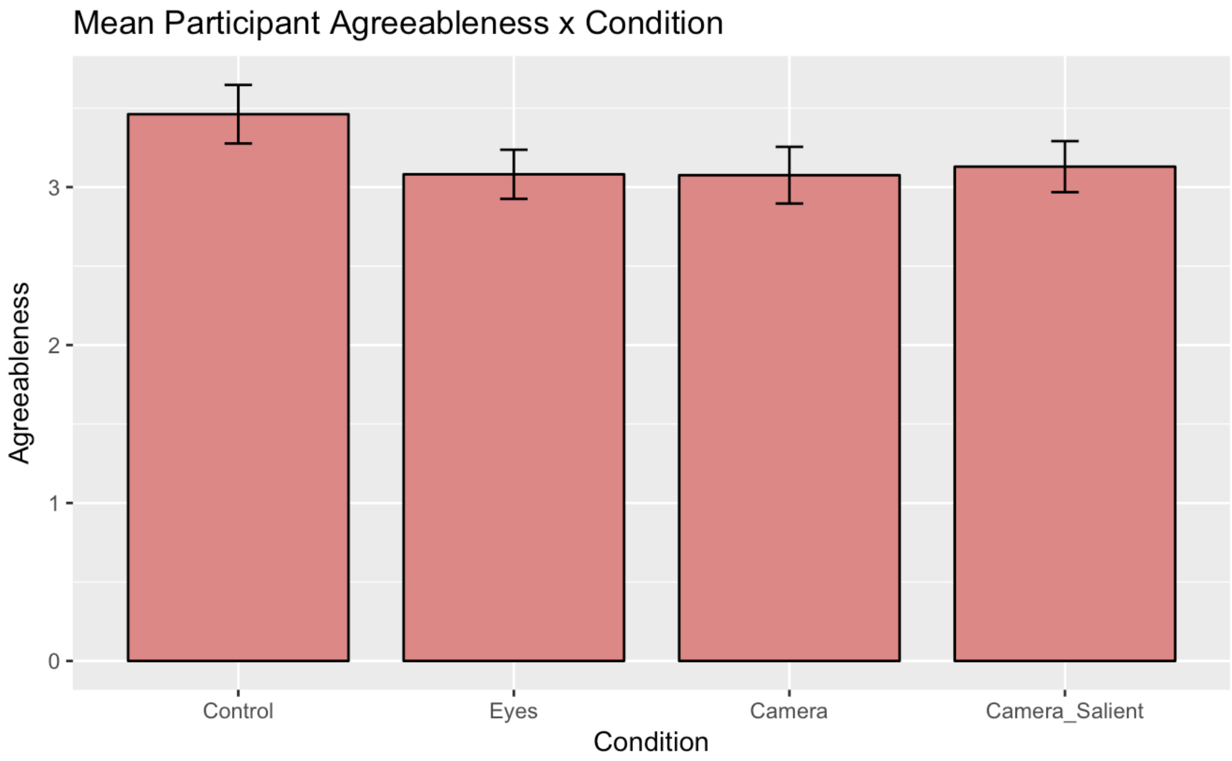


*Figure 2.4. Study 1. Mean HEXACO PI-R Agreeableness Scores by Condition. Error bars represent the 95% Confidence Interval.*

Participants reported higher levels of conscientiousness in the control condition (*M* = 3.80, *SD* =0.56) than in the eyes condition (*M* =3.63, *SD* = 0.53) but lower levels in the control condition when compared with the Camera (*M* = 3.85, *SD* = 0.50) and Camera Salient (*M* = 3.94, *SD* = 0.55) conditions. A one way ANOVA suggested that the differences were significant, $F(3, 203) = 3.17, p = .03, \omega^2 = 0.03$. Figure 2.5 shows mean conscientiousness scores by condition.

However, further exploration of conscientiousness scores through a series of Welsh two sample t-tests showed that the differences between each condition were not significantly different. Not between the Control condition and Eyes, $t(100.09) = 1.64, p = 0.10, r = .16$. Nor in the Control condition compared with either Camera, $t(96.85) = -0.42, p = 0.68, r = .04$ or in the Control compared with the Camera Salient condition, $t(94.99) = -1.21, p = 0.23, r = .12$.

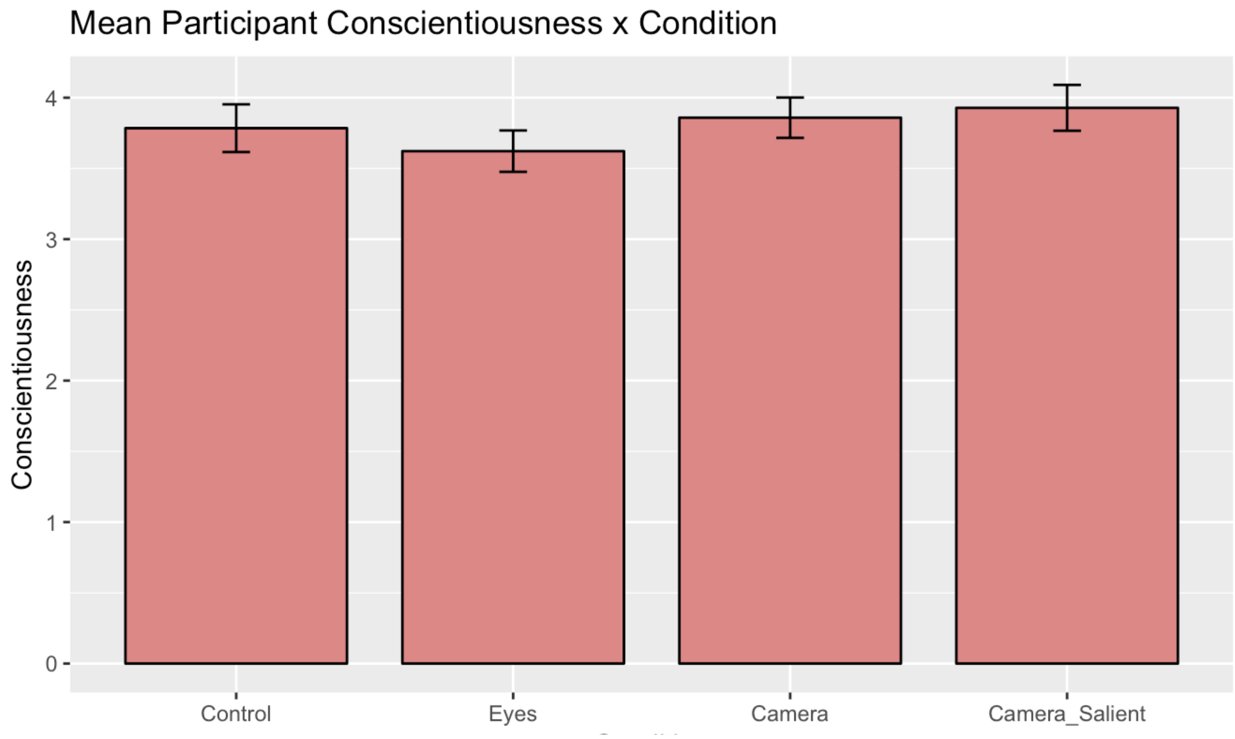


*Figure 2.5. Study 1. Mean HEXACO-PI-R Conscientiousness Scores by Condition. Error bars represent the 95% Confidence Interval.*

### 2.3.1. Zurich Prosocial Game – Interaction Effects Analysis

Table 2.1. showed that in the ZPG, participants helped no more in the eyes ($M = 0.55$, $SD = 0.23$), camera ($M = 0.53$, $SD = 0.22$) or camera salient ($M = 0.56$, $SD = 0.24$) conditions than they did in the control ($M = 0.50$, $SD = 0.24$), $F(3, 203) = 0.55$, $p = 0.65$, $\omega^2 = -0.14$. Table 2.2. presents a full hierarchical generalised linear model (GLM) in four steps exploring interaction effects in more detail, given the significant interaction between condition and HEXACO PI-R dimensions found in our first analysis. In Step 1 we test the main effect of condition on frequency of helping with control condition as the comparison, showing no main effect of condition on frequency of helping in the ZPG. The adjusted $R^2$ for model 1 shows that condition explained <1% of the variance in participant helping behaviours. In Step 2 we add in Snyder Self-Monitoring (SM) Scores, finding no significant main effect of condition on SM Scores. In Step 3 we look for interaction effects between surveillance condition and SM Score on helping behaviour, with still just 1% of the variance in helping explained. In Step 4 we add our exploratory analysis looking for interactions between condition and the six dimensions of the HEXACO PI-R as predictors of helping behaviour. Not surprisingly, adding the personality factors did improve the model, with 14% of the variance in helping behaviours explained. However, none of the interactions between condition and HEXACO PI-R nor those between condition and SM Score proved significant. Comparing the models in Step 4 we can see that the changes in $R^2$ were not significant.

Additionally there were no interaction effects between Snyder Self-Monitoring Scores and condition on the frequency of helping, $F(3, 199)$, = 0.14, $p = 0.94$, $\omega^2 = -0.14$. Nor was there any interaction between condition and any HEXACO PI-R personality dimensions on frequency of helping as shown in Step 4, $F(31, 175)$, = 1.04, $p = 0.43$.

The only main effect was that of emotionality $B = 0.14$, $SE\ B = 0.06$, $t = 2.18$, $p = 0.03$, (see the scatterplot in Figure 2.6. below). As can be seen, emotionality predicted more frequent helping behaviours across all conditions.

**Table 2.2. Hierarchical GLM Examining Helping Behaviour in the ZPG by surveillance condition, SM Score & HEXACO Personality Index Measures**

| | $\Delta R^2$ (Adjusted) | *B* | *SE B* | *t value* | *P* |
|---|---|---|---|---|---|
| Step 1 (Model 1), F(3, 203)= 0.55, p = 0.65 | 0.008 | | | | 0.65 |
| Constant | | 0.50 | 0.03 | | |
| Condition: Eyes | | 0.05 | 0.05 | 1.05 | 0.30 |
| Condition: Camera | | 0.02 | 0.05 | 0.48 | 0.63 |
| Condition: Camera Salient | | 0.05 | 0.05 | 1.11 | 0.27 |
| Step 2 (Model 2), F(4, 202) = 0.51, p = 0.73 | 0.002 | | | | 0.73 |
| Constant | | 0.48 | 0.06 | 8.28 | |
| Condition: Eyes | | 0.05 | 0.05 | 1.05 | 0.30 |
| Condition: Camera | | 0.02 | 0.05 | 0.53 | 0.60 |
| Condition: Camera Salient | | 0.05 | 0.05 | 1.16 | 0.25 |
| SM Score | | 0.002 | 0.004 | 0.62 | 0.54 |
| Step 3 (Model 3), F(7, 199), = 0.35, p = 0.93 | 0.002 | | | | |
| Constant | | 0.47 | 0.10 | | |
| Condition: Eyes | | 0.07 | 0.15 | 0.48 | 0.63 |
| Condition: Camera | | 0.06 | 0.14 | 0.47 | 0.64 |
| Condition: Camera Salient | | 0.02 | 0.14 | 0.15 | 0.88 |
| SM Score | | 0.003 | 0.01 | 0.37 | 0.71 |
| Condition Eyes: SM Score | | -0.002 | 0.01 | -0.17 | 0.86 |
| Condition Camera: SM Score | | -0.004 | 0.011 | -0.32 | 0.75 |
| Condition Camera Salient: SM Score | | 0.003 | 0.01 | 0.27 | 0.78 |

| …cont'd. | $\Delta R^2$ (Adjusted) | *B* | *SE B* | *t value* | *P* |
|---|---|---|---|---|---|
| Step 4 (Model 4), F(31, 175), = 1.04, p = 0.43 | 0.143 | | | | 0.43 |
| Constant | | 0.32 | 0.53 | 0.59 | |
| Condition: Eyes | | 0.55 | 0.62 | 0.88 | 0.38 |
| Condition: Camera | | 0.15 | 0.65 | 0.23 | 0.82 |
| Condition: Camera Salient | | -0.12 | 0.66 | -0.18 | 0.85 |
| SM Score | | 0.01 | 0.01 | 0.98 | 0.32 |
| Honesty-Humility | | -0.01 | 0.06 | -0.12 | 0.90 |
| Emotionality | | 0.14 | 0.06 | 2.18 | **0.03*** |
| Extraversion | | -0.02 | 0.07 | -0.33 | 0.74 |
| Agreeableness | | -0.05 | 0.06 | -0.93 | 0.36 |
| Conscientiousness | | 0.02 | 0.06 | 0.35 | 0.73 |
| Openness | | -0.02 | 0.06 | -0.30 | 0.76 |
| Condition Eyes: SM Score | | -0.12 | 0.01 | -0.90 | 0.90 |
| Condition Camera: SM Score | | -0.01 | 0.01 | -0.38 | 0.70 |
| Condition Camera Salient: SM Score | | -0.003 | 0.01 | -0.29 | 0.77 |
| Condition Eyes: Honesty-Humility | | -0.04 | 0.08 | -0.05 | 0.96 |
| Condition Camera: Honesty-Humility | | 0.13 | 0.09 | 1.43 | 0.16 |
| Condition Camera Salient: Honesty-Humility | | -0.01 | 0.09 | -0.07 | 0.95 |
| Condition Eyes: Emotionality | | -0.05 | 0.08 | -0.67 | 0.50 |
| Condition Camera: Emotionality | | -0.09 | 0.07 | -1.22 | 0.22 |
| Condition Camera Salient: Emotionality | | -0.01 | 0.09 | -0.14 | 0.89 |
| Condition Eyes: Extraversion | | 0.09 | 0.10 | 0.98 | 0.33 |

| …Step 4 cont'd... | $\Delta R^2$ (Adjusted) | *B* | *SE B* | *t value* | *P* |
|---|---|---|---|---|---|
| Condition Camera: Extraversion | | 0.07 | 0.10 | 0.70 | 0.48 |
| Condition Camera Salient: Extraversion | | 0.04 | 0.10 | 0.38 | 0.71 |
| Condition Eyes: Agreeableness | | -0.02 | 0.09 | -0.20 | 0.84 |
| Condition Camera: Agreeableness | | -0.06 | 0.08 | -0.74 | 0.47 |
| Condition Camera Salient: Agreeableness | | 0.03 | 0.09 | 0.35 | 0.73 |
| Condition Eyes: Conscientiousness | | 0.11 | 0.09 | -1.17 | 0.25 |
| Condition Camera: Conscientiousness | | -0.09 | 0.09 | -0.93 | 0.35 |
| Condition Camera Salient: Conscientiousness | | -0.003 | 0.09 | -0.04 | 0.97 |
| Condition Eyes: Openness | | -0.04 | 0.07 | -0.58 | 0.56 |
| Condition Camera: Openness | | -0.01 | 0.08 | -0.15 | 0.88 |
| Condition Camera Salient: Openness | | 0.002 | 0.09 | 0.03 | 0.98 |
| ANOVA Models 1-4 | | | | | |
| Model 1: Donated by Condition | (203) RSS = 10.84 | | | | |
| Model 2 vs Model 1 | F(1, 202) = 0.39, *p* = 0.53, RSS = 10.82 | | | | |
| Model 3 vs Model 2 | F(3, 199) = 0.15, *p* = 0.93, RSS = 10.80 | | | | |
| Model 4 vs Model 3 | F(24, 175) = 1.23, *p* = 0.22, RSS = 9.24 | | | | |

* = significant at alpha < .05, ** = significance < .01, *** = significance < .001

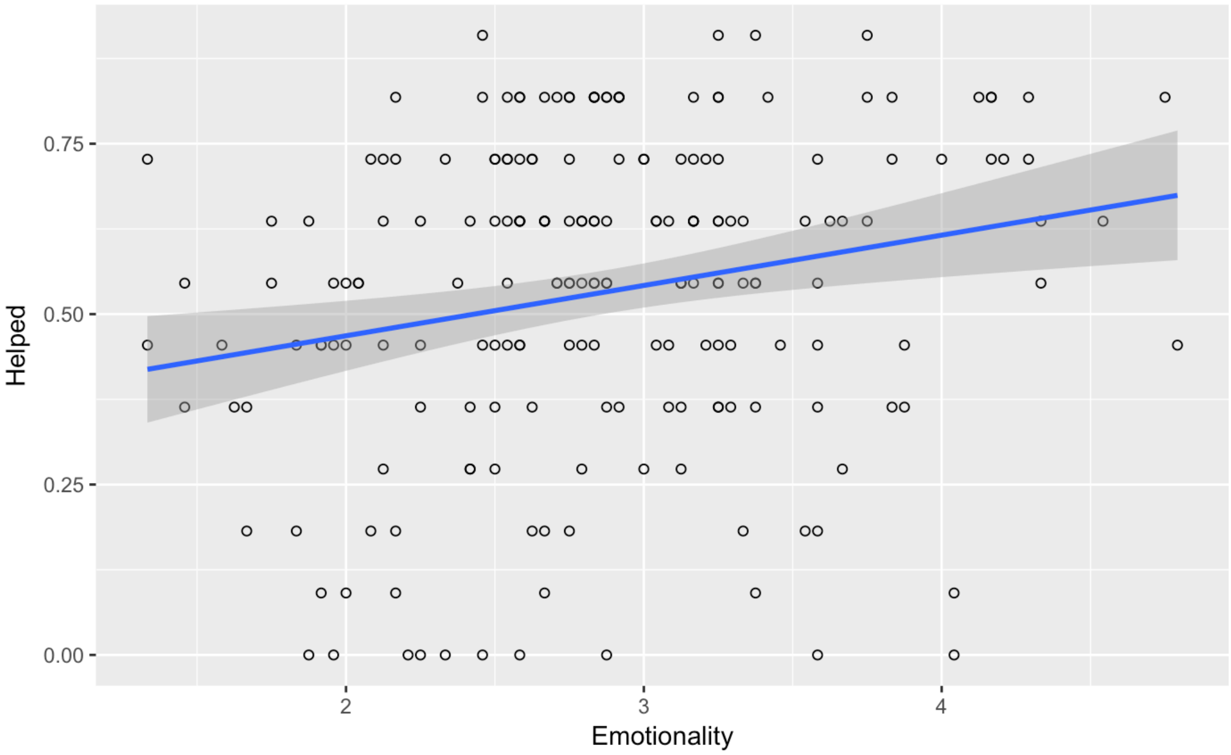


*Figure 2.6. Study 1. The influence of emotionality on helping behaviours in the ZPG. Greyed area represents the 95% Confidence Interval*

### 2.3.2. Are the Zurich Prosocial Game and Dictator Game Measuring the Same Behaviour?

A Pearson's correlation test between frequency of helping behaviours in the ZPG and generosity of donations in the Dictator Game found no significant relationship between the ZPG and Dictator Games as measures of prosociality, $R^2$ =0.0081, $p$ =.20. Our results suggest that the ZPG and Dictator Game may not be measuring the same phenomena. Only 0.81% of the variance in donations is accounted for by frequency of helping behaviours suggesting that our two tests were measuring different things and not a common underlying prosociality.

### 2.3.3. Dictator Game Results – Interaction Effects Analysis

As we have seen already from Table 2.1. above, there was no overall effect of condition on donation in the Dictator Game *F* (3, 203) = 0.20, $p = 0.90$, $\omega^2 = 0.01$. Table 2.3. presents the hierarchical GLM in full.

Step 1 shows that there was no main effect of condition on donation and that condition (Model 1) explained less than 1% of the variance in donations. Step 2 shows that there was no significant effect of condition on donation when taking into account self-monitoring scores, and that even after adding self-monitoring scores to look for interaction effects between SM Score and Condition on donations (Step 3) the model still explained only 1% of the variance. Steps 2 and 3 showed there were no siginificant interaction effects of self-monitoring scores with condition, nor any interactions between the HEXACO PI-R personality components and condition on donations.

The only significant main effect was that of the Honesty-Humilty personality dimension on donations, $B = 1.83$, *SE B* = 0.81, t = 2.25, $p < 0.05$. More specifically, higher HH scores correlated with decreased donations. However, a Pearson's correlation test for this effect was not significant *t* (197) = -1.06, *r* = - .08 p = .28, *95% CI* [-.21, .06]. Even when combined with surveillance condition, self-monitoring scores and all HEXACO PI-R personality dimensions in our Model (Model 4) it still explained just 17% of the variance in donations. Comparing the models we can see that the changes in $R^2$ were not significant.

**Table 2.3. Hierarchical Linear Model Examining Dictator Game Donations by Condition, SM Score & HEXACO Personality Index Measures**

| | $\Delta R^2$ | *B* | *SE B* | *t value* | *P* |
|---|---|---|---|---|---|
| Step 1 (Model 1), F(3, 203)= 0.20, p = 0.90 | 0.003 | | | | 0.90 |
| Constant | | 5.16 | 0.44 | | |
| Condition: Eyes | | -0.22 | 0.59 | -0.36 | 0.72 |
| Condition: Camera | | -0.45 | 0.60 | -0.74 | 0.46 |
| Condition: Camera Salient | | -0.12 | 0.62 | -0.20 | 0.84 |
| Step 2 (Model 2), F(4, 202) = 0.22, p = 0.93 | 0.001 | | | | 0.93 |
| Constant | | 5.50 | 0.76 | 7.26 | |
| Condition: Eyes | | -0.22 | 0.59 | -0.37 | 0.72 |
| Condition: Camera | | -0.48 | 0.62 | -0.24 | 0.81 |
| Condition: Camera Salient | | -0.15 | 0.62 | -0.24 | 0.81 |
| SM Score | | -0.03 | 0.05 | -0.54 | 0.59 |
| Step 3 (Model 3), F(7, 199) = 0.54, p = 0.80 | 0.014 | | | | |
| Constant | | 5.78 | 1.37 | 4.24 | |
| Condition: Eyes | | -0.96 | 1.94 | -0.50 | 0.62 |
| Condition: Camera | | 0.51 | 1.78 | 0.29 | 0.77 |
| Condition: Camera Salient | | -1.70 | 1.81 | -0.94 | 0.35 |
| SM Score | | -0.05 | 0.11 | -0.40 | 0.69 |
| Condition Eyes: SM Score | | 0.06 | 0.16 | 0.40 | 0.69 |
| Condition Camera: SM Score | | -0.09 | 0.15 | -0.64 | 0.52 |
| Condition Camera Salient: SM Score | | 0.14 | 0.15 | 0.94 | 0.25 |

2. Study 1 Perceived Surveillance & Prosociality

| …cont'd. | $\Delta R^2$ (Adjusted) | *B* | *SE B* | *t value* | *P* |
|---|---|---|---|---|---|
| Step 4 (Model 4), F(31, 175), = 1.30, p = 0.15 | 0.168 | | | | 0.15 |
| Constant | | -3.52 | 6.87 | -0.51 | 0.65 |
| Condition: Eyes | | 3.64 | 8.01 | 0.46 | 0.65 |
| Condition: Camera | | 3.05 | 8.45 | 0.36 | 0.72 |
| Condition: Camera Salient | | 7.77 | 8.56 | 0.91 | 0.37 |
| SM Score | | -0.03 | 0.13 | -0.26 | 0.80 |
| Honesty-Humility | | 1.83 | 0.81 | 2.25 | **0.03*** |
| Emotionality | | -0.78 | 0.80 | -0.97 | 0.34 |
| Extraversion | | 0.27 | 0.96 | 0.28 | 0.78 |
| Agreeableness | | 0.38 | -.76 | 0.51 | 0.61 |
| Conscientiousness | | 0.40 | 0.84 | 0.48 | 0.63 |
| Openness | | 0.14 | 0.79 | 0.18 | 0.86 |
| Condition Eyes: SM Score | | 0.10 | 0.18 | 0.53 | 0.60 |
| Condition Camera: SM Score | | -0.08 | 0.18 | -0.46 | 0.65 |
| Condition Camera Salient: SM Score | | 0.17 | 0.17 | 1.00 | 0.32 |
| Condition Eyes: Honesty-Humility | | -1.45 | 1.08 | -1.35 | 0.18 |
| Condition Camera: Honesty-Humility | | -0.41 | 1.14 | -0.36 | 0.72 |
| Condition Camera Salient: Honesty-Humility | | -0.89 | 1.12 | -0.79 | 0.42 |
| Condition Eyes: Emotionality | | 1.23 | 1.04 | 1.19 | 0.24 |
| Condition Camera: Emotionality | | 1.08 | 0.94 | 1.15 | 0.25 |
| Condition Camera Salient: Emotionality | | 0.80 | 1.17 | 0.69 | 0.49 |
| Condition Eyes: Extraversion | | -1.09 | 1.24 | -0.88 | 0.38 |

| …Step 4 cont'd... | $\Delta R^2$ (Adjusted) | *B* | *SE B* | *t value* | *P* |
|---|---|---|---|---|---|
| Condition Camera: Extraversion | | -0.49 | 1.31 | -0.37 | 0.71 |
| Condition Camera Salient: Extraversion | | -0.66 | 1.32 | -0.50 | 0.62 |
| Condition Eyes: Agreeableness | | 1.56 | 1.11 | 1.40 | 0.16 |
| Condition Camera: Agreeableness | | -0.15 | 1.08 | -0.14 | 0.89 |
| Condition Camera Salient: Agreeableness | | -0.27 | 1.15 | -0.23 | 0.82 |
| Condition Eyes: Conscientiousness | | -1.36 | 1.20 | -1.13 | 0.26 |
| Condition Camera: Conscientiousness | | -1.11 | 1.21 | -0.91 | 0.36 |
| Condition Camera Salient: Conscientiousness | | -1.16 | 1.18 | -0.99 | 0.33 |
| Condition Eyes: Openness | | 0.39 | 0.95 | 0.41 | 0.68 |
| Condition Camera: Openness | | 0.82 | 1.00 | 0.82 | 0.41 |
| Condition Camera Salient: Openness | | -0.19 | 1.11 | -0.17 | 0.86 |
| ANOVA Models 1-4 | | | | | |
| Model 1: Donated by Condition | (203) RSS = 1886.2 | | | | |
| Model 2 vs Model 1 | F(1, 202) = 0.31, *p* = 0.58, RSS = 1883.5 | | | | |
| Model 3 vs Model 2 | F(3, 199) = 1.03, *p* = 0.38, RSS = 1856.3 | | | | |
| Model 4 vs Model 3 | F(24, 175) = 1.51, *p* = 0.07, RSS = 1538.4 | | | | |

* = significant at alpha < .05, ** = significance < .01, *** = significance < .001

### 2.3.4. Snyder Self-Monitoring Scores & Prosocial Behaviours

We hypothesized, after the findings of Van Bommel et al. (2012), that an individual might feel more watched in the surveillance condition than in the control condition and thus they might report higher scores when completing the Snyder Self-Monitoring Scale questionnaire in the Eyes, Camera, and Camera (Salient) condition than in the control. However, an ANOVA examining SM scores by condition showed this was not the case. Scores did not significantly differ by condition $F(3, 203) = 0.94$, $p = 0.42$, $\omega^2 < 0.001$ (see Table 2.1. above and Figure 2.7. below).

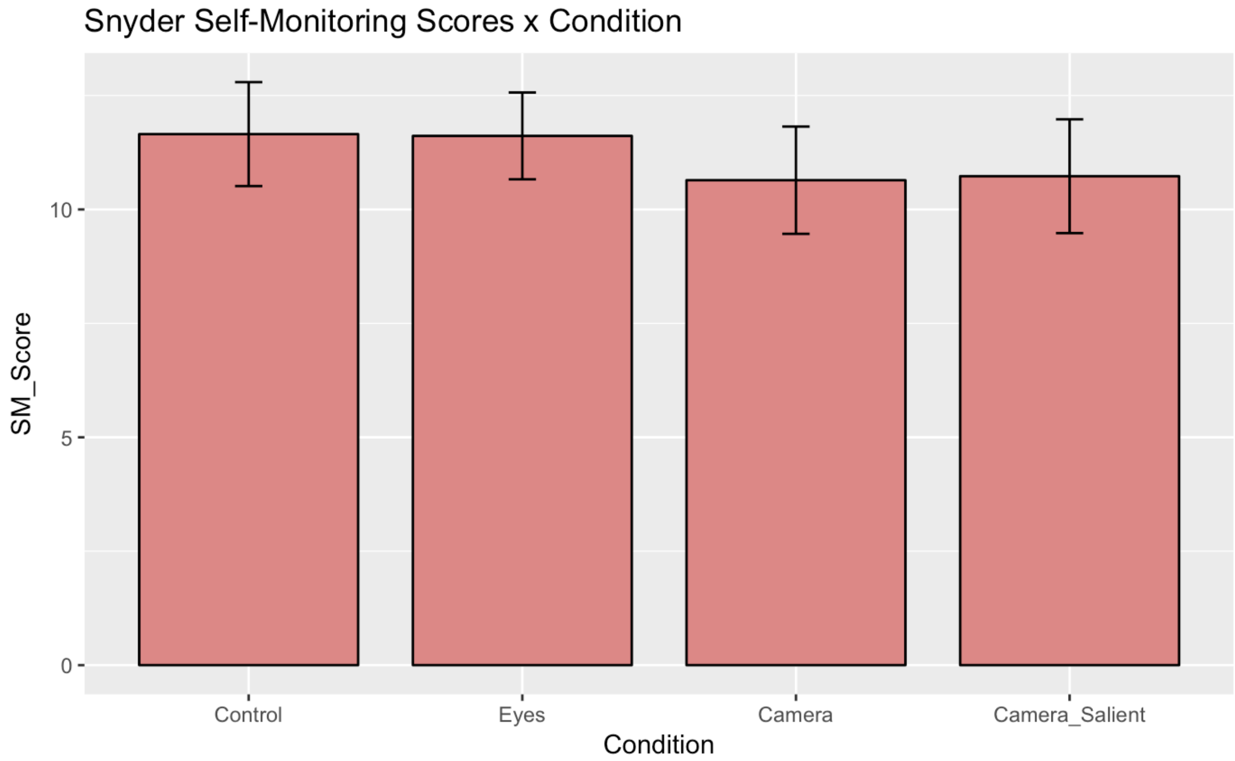


*Figure 2.7. Study 1. Mean Snyder Self-Monitoring Scores by Condition*

## 2.4. Discussion

This was a large-scale and carefully controlled experiment. Our failure to replicate the effect of watching eyes on donations in the Dictator Game, or to find any effect of eye cues on how often participants helped others in the ZPG adds to the corpus of literature that is beginning to question whether 'watching eyes' influence generosity specifically, or indeed whether they influence behaviour at all (Northover, Pedersen, Cohen, & Andrews, 2017a, 2017b). There were no interaction effects of self-monitoring scores or the HEXACO personality dimensions and task conditions on either donations or helping behaviours. This failure to find any eye cue effects in our study suggests that perhaps previous effects are the result of publication bias or at least are highly context dependent.

If we see the self-monitoring scores as a manipulation check – to see if the presence of eye cues actually made people feel more watched, our study can be said to have failed to make people feel watched in the surveillance cue conditions and thus it is not surprising that we found no effect on helping behaviours or donations. However, our eye images were taken from Nettle et al.'s (2013) study which did find an effect of eye cues on generosity in the dictator game. Consequently, we would need to explain why these eye cues had an effect in one study and not in another. The most parsimonious explanation is that surveillance had no effect on generosity, and previous findings have been false positives. That said, we cannot be certain that self-consciousness is an effective proxy measure of how watched someone feels. We measured public self-consciousness not as a manipulation check, but as an additional exploratory analysis only. It is premature to write off our examination of the watching eyes effect on prosocial behaviour based on the absence of any effect of surveillance cues on self-consciousness until future research demonstrates it to be an effective manipulation check.

The presence of a webcam, regardless of whether it was switched off or on, also had no effect on how frequently participants helped others. With regard to this issue, it is possible that Van Bommel et al.'s study (2012) had greater ecological validity. Our participants were completing a computer-based task whereas the aforementioned design had people present in a room in a simulated scenario where participants did not know the experiment had begun. Nevertheless, if taken at face value, our findings undermine the claims of advocates for widespread surveillance and the philosophical assertions of those such as Bentham, who argue that causing a person to feel watched increases prosocial behaviour (Bentham, 1791). The results also undermine the concerns of privacy campaigners who argue that feeling watched causes chilling effects (self-censorship), since feeling watched had no effect in our study (Hermstruwer & Dickert, 2017; Penney, 2016).

To this end our results align with those of Buchanan et al. (2017) who found no effect of cameras on Dictator Game decisions – even when participants knew that their decisions were being filmed and were being screened live on a monitor to a group of other participants/researchers. Our study upholds the previous findings of these authors and suggests that the mere presence of a camera has little effect on prosocial behaviour.

We predicted that participants who identified themselves as high self-monitors – i.e. individuals who are more prone to adapting their behaviour to different contexts and audiences - would be more responsive to surveillance cues. We expected that they would help and donate more when eye cues or a camera were present in comparison with how frequently they helped and how much they donated in the control condition. By contrast, low-self monitors, who are generally consistent in their behaviour regardless of who or what is present, were predicted to be largely unresponsive to surveillance cues. In our experiment, there was no evidence that either group was affected by surveillance cues, with no effect of surveillance condition on self-monitoring scores and no interaction effect condition by self-monitoring on helping behaviours in either the ZPG or donations in the dictator game.

We did find a main effect of surveillance condition on emotionality, agreeableness and conscientiousness (Table 1, Figures 3, 4, and 5). None of these main effects reached the threshold of $\omega^2 = .01$ generally regarded as a small effect size (Kirk, 1996). *t*-tests found that conscientiousness scores in each surveillance cue condition where not significantly different from those in the no-surveillance cue control, while agreeableness scores fell in the opposite direction to that predicted – those in the control were more agreeable than those in the surveillance cue condition. This, then, leaves just the main effect of condition on emotionality to be discussed.

Before we proceed, it is important to note that HEXACO PI-R personality dimensions were included in our analysis principally as an exploratory trait measure and covariate. We nevertheless found a main effect of surveillance condition on some personality dimensions. The simplest explanation for this is that the randomization procedure did not work. If personality facets are traits then our attempts at randomization across conditions failed and no further exploration of the effects of surveillance cue on HEXACO personality dimensions is necessary.

However, there are well-documented self-report vs observer biases in personality assessments that might mean that our personality dimensions were influenced by surveillance cues. Individuals tend to report higher scores on traits related to self-efficacy than those who observe and report on them do (Ashton & Lee, 2010). There is also evidence for automatic, or unconscious adjustment in self-presentation when participants' responses in personality surveys are in public rather than in private (Paulhus & Williams, 2002). Evaluative biases in self-reporting are also distinguished by those that are agentive (getting ahead) vs communal and moralistic (getting along). Individuals are said to emphasise the communal and moralistic in public but the agentive in private (Paulhus & Williams, 2002). These self-presentation biases provide a possible explanation for why the HEXACO scores for conscientiousness, emotionality and agreeableness differ in the control when compared with the surveillance cue condition – perhaps watching eyes created evaluation apprehension as if the participant was being watched - they are therefore explored further.

Our *t*-tests showed that emotionality scores in the control were significantly lower than those reported in each of the surveillance cue conditions with an effect size of $r$ =.29 in the Eyes, $r$ = .28 in the camera, and $r$ = . 25 in the camera salient conditions These are above a small effect as according to Pearson's heuristic which sets $r$ = .1 as a small effect and $r$ = .3 as a medium effect and $r$ >.5 as a large effect (Cohen, 1992). At $r$ = . 25, the lowest of the three effect sizes, for every 100 people in each condition 8 more would report higher scores in the camera than in the control. As a more simple measure, participants reported emotionality scores that were 13% lower in the control than in the eyes condition, 8% lower in the control than the camera condition and 9% lower in the control than in the camera salient condition. Emotionality is a personality factor defined by traits that facilitate altruism *directly* and *indirectly*: *directly* via empathy and concern for others and the tendency to become emotionally attached to others; *indirectly,* via the desire to avoid harm and willingness to seek help. It is possible that the presence of a surveillance cue caused participants to be primed towards a more altruistic disposition, with perhaps the tasks themselves being insufficiently sensitive or ecologically valid for this disposition to translate into behaviour. Thinking about this finding from an applied perspective, it might be more meaningful on a large scale. A 13% increase in the predisposition towards altruism in the presence of eyes cues or an 8-9% increase in the presence of cameras, might translate into a smaller percentage change in prosocial behaviour than would be sufficent to be discerned in a relatively small-scale lab experiment. Nevertheless, such an effect might still have important public policy implications where improving compliance with desired behaviours by even just 1% can be important, saving lives and money.

That said, however there was no interaction effect between surveillance condition and emotionality on helping behaviours or donations, although, as we caveat, it might be observed that the fact that there was no significant correlation between helping and donations in our two prosocial games suggests that the lack of an interaction effect may reflect a problem with the dictator game or the ZPG as a measure to test prosociality, not just their sensitivity. On the other hand the the main effect of emotionality on helping in the ZPG is as would be expected: more emotional participants help more. So we have slightly contradictory indicators. That the ZPG and dictator game outcomes do not correlate suggests they are not testing the same thing. That more prosocial participants help more suggests the ZPG does test prosociality. It could be that the dictator game doesn't test generosity. Or it might be that surveillance cues have variable effects on different types of generosity as discussed in the next chapter.

Finally, the variable direction of the effects might be explained by that fact that participants do not react uniformly to the switch from private to public context. For example, people can present themselves to embody intimidation, supplication, ingratiation, self-promotion or exemplification (Paulhus & Williams, 2002). Some personality types might show consistent biases in their self-presentation (Ashton & Lee, 2010) but we had no measure to detect these biases since we analysed by aggregated individual personality dimensions only. It is plausible, then, that the effect of surveillance cues would be markedly different on different individuals. If so, a within-subjects design might be better able to test this. As far as the present study goes, we can say only that the findings suggest that the surveillance cues might have inclined participants towards altruism, but that this did not translate into more altruistic behaviour.

## 2.5. Limitations

There are three principal limitations to note within this study: the possibility of order and framing effects, limitations in the design of the Dictator game, and limitations in the statistical power of the experiment.

Participants in all groups always completed the ZPG prior to the Dictator Game, and therefore participants could have been influenced in their Dictator Game decisions by the framing of their immediately preceding experience in completing the ZPG or perhaps by fatigue. However, studies which have specifically aimed to create deliberate and egregious framing effects to test their influence - such as by manipulating the names of the Dictator Game with the morally valanced terms ‘giving or taking games’ and similarly titling actions within the task, have found that framing has no effect. Following two such experimental manipulations Dreber, Ellingsen, Johannesson, and Rand (2013) were moved to ask ‘Do People Care About Context?’. The answer from their work was a clear no. So, while it is right to note the risk of order effects, given evidence that it makes little difference, perhaps we should not detain ourselves overly with the concern that playing the ZPG first would have influenced the Dictator Game decisions enough to confound the results.

Secondly, in the debrief, one participant suggested that the Dictator Game did not necessarily test generosity. Participants knew that they would receive an allocation from a previous ‘random participant’ on completion of the experiment. They also knew that that their own allocation would go to a future ‘random participant’. Therefore, the experiment might cause a participant to judge what they thought others would give not to maximise their own gain, but to achieve equilibrium in donations and thus a ‘fair’ outcome. If they expect to receive the same amount from a random previous participant as they put in the envelope for a future participant they cannot be judged to have acted unfairly or generously irrespective of whether they kept all or gave all. This is an important critique. With hindsight, participants should not have been told they were receiving an envelope from a previous participant. They were told the truth in order to minimise deception, but in this instance, this small deception would have been justified. To avoid such a deception, we could have divided participants into receiver and dictator roles, but this would have required a much larger sample and would have been costlier since we would still have had to pay participants a 'show-up fee' in addition to any possible money they might have donated to them in the task. Nevertheless, although a few participants clearly did think deeply enough about it to consider how to achieve equilibrium, the majority talked afterwards during the debrief in terms of generosity and selfishness. Additionally, since participants have no idea what the equilibrium will be the experiment still tests generosity, just less precisely than intended. Participants who give do so knowing they might end up receiving less, participants who don't do so know they might profit from the generosity of others. Thus, while there is room for improvement in our methods, we still believe this experiment was a valid test of the effect of surveillance cues on prosocial behaviour.

A more serious limitation surrounds the power of our study. Our sample size was 207, Control *N* = 49, Eyes *N* = 57, Camera *N* = 53, Camera Salient *N* = 49 and was broadly representative of sample size in similar studies at the time the research was designed. For example: Tane and Takezawa (2011), N= 40 participants; Oda, Niwa, Honma, Hiraishi (2011), where N=62; Kitayama, Snibbe, Markus, and Suzuki (2004), N=52, 93 & 65; or G. Hoffman et al. (2015) where N=60. A G*Power analysis (Faul, Erdfelder, Lang, & Buchner, 2007) suggests the sample size necessary to detect a small effect (F=0.1) in a one-way ANOVA comparing participants mean helping in the ZPG or mean donations in the dictator game by Condition (*df*=3), a power of .8 and α error probability of 0.05 required a sample size of 1096. Therefore, we would not have had the power to detect small effects which would be meaningful if considered in terms of our broader research interest – even a 1%

improvement in prosocial behaviour in, say, a city-wide surveillance initiative might be sufficient to justify it proceeding.

## 2.6. Conclusion

Overall, our null findings align with recent meta-analyses and reviews conducted by Northover et al. (2017a) in finding no significant influence of eye cues on generosity. The results of this chapter raise the question as to whether there is a watching eyes effect at all. As mentioned previously, our findings undermine both Bentham's conceptual claim that a panoptic society promotes prosocial behaviour, and the claims of Foucault that perceived surveillance leads to reflexive self-censorship. Our findings also weaken the case for modern proponents of surveillance interventions in public policy, while simultaneously disarming some of the arguments made by critics of state and commercial surveillance by suggesting that surveillance alone is not a deterrent – perhaps it must have consequences to exert influence over behaviour.

2. Study 1 Perceived Surveillance & Prosociality

# Chapter 3

An edited version of this chapter has been accepted for publication by the journal *Evolution and Human Behavior* under the following title:

**Dear, K. P.,** Dutton, K. J., & Fox, E. (in press). Do 'Watching Eyes' Influence Antisocial Behaviour? A Systematic Review & Meta-Analysis.

# 3

# A Systematic Review & Meta-analysis of the ‘Watching Eyes Effect’ on Antisocial Behaviour

> *“But fear Allah and know that Allah sees well what ye do.”*
>
> Surah Al-Baqarah, 233

Chapter 1 presented evidence suggesting why there might be an effect of eye cues on behaviour. Chapter 2 found no effect of the presence of eye cues or cameras on prosociality. Chapter 3 systematically reviews the eye cue literature noting failed replication attempts, and two meta-analyses which raise doubts about the reproducibility of the *watching eyes* effect on generosity. However, it also highlights that much of the wider evidence on eye cues has still not been systematically reviewed, notably that which is most relevant to its practical application: the effect of eye cues on antisocial behaviour. Given the evidence of humans' heightened sensitivity to threat and negative information, it is hypothesized that the watching eyes effect would be more consistent on antisocial behaviour. In a meta-analysis of 15 experiments from 13 research papers we find a reduction in the risk of antisocial behaviour of 35% when eye cues are present. By contrast, reviews suggest that CCTV cameras reduce crime by only 16%. Our meta-analytic evidence for a *watching eyes* effect on antisocial behaviour is sufficient to justify the use of eye cues in the very low-cost and potentially high-impact real-world interventions that are proliferating in public policy, particularly in the UK.

# 3. A Systematic Review & Meta-analysis of the 'Watching Eyes Effect' on Antisocial Behaviour

## 3.1. Introduction

Chapter 1 selected the watching eyes effect as a research focus for this thesis, a way in to understanding whether there is an effect of surveillance that is reflexive, automatic, or unconscious – something akin to Bentham's idea of the way the panopticon would improve behaviour, Foucault's 'reflexive control', privacy campaigners 'chilling effects' or humanitarian's Ambient Protective Effect. Chapter 2 dived headlong into this research with Study 1 seeking to test in a large-scale and carefully controlled experiment whether watching eyes increased prosocial behaviour in two tasks, extending the paradigm to see if a camera might have similar effects, and adding self-monitoring and personality profiling measures as exploratory possible covariates. That Study 1 found no main effect of surveillance cues prompted the systematic review and meta-analysis that follows

Much of the research on the influence of pictures of eyes on behaviour has focused on prosocial behaviour. For instance, a picture of human eyes added to charity donation buckets in supermarkets increased donations by 48% in comparison with a control condition (Powell, Roberts & Nettle 2012).

In addition to field experiments, laboratory studies have also addressed the question of whether eye cues increase generosity, utilizing the Dictator Games (Nettle, Harper, et al., 2013) and similar economic games. Two of the earliest studies in the field serve as examples. In 2005, Haley & Fessler found subtle watching eye cues increased generosity in the Dictator game by 31%, and when the eye cues were more clearly visible by 55%. A similar eye cues & economic game design run by Burnham & Hare in 2007 suggested that the presence of eyes increased altruistic contributions in an economic simulation by 29% in comparison with a control.

# 3. A Systematic Review & Meta-analysis of the 'Watching Eyes Effect' on Antisocial Behaviour

More recent research has challenged earlier findings on the effect of eyes cues on prosocial behaviour. A series of replication attempts and new experiments failed to find any consistent evidence for an eye cue effect (Beyfus et al., 2016; Brudermann, Bartel, Fenzl, & Seebauer, 2015; Bush, Erlich, Prather, & Zeira, 2016; Cai, Huang, Wu, & Kou, 2015; Carbon & Hesslinger, 2011; Fehr & Schneider, 2010; Fujii, Takagishi, Koizumi, & Okada, 2015; Gaiani, Rose, & Roberts, 2014; Golja, 2013; Huang, Liu, Zheng, Tan, & Zhao, 2015; Jackson, 2015; Jolij & de Haan, 2014; Kuliga, Tanja-Dijkstra, & Verhoeven, 2011; Lamba & Mace, 2010; Matland & Murray, 2015; Matsugasaki, Tsukamoto, & Ohtsubo, 2015; Northover, Pedersen, Cohen, & Andrews, 2017; Palomäki, Modic, & Yan, 2015; Raihani & Bshary, 2012; Rose, Gaiani, & Roberts, 2014; Sparks, 2010; Sparks & Barclay, 2015; Stella et al., 2013; Tane & Takezawa, 2011; Vogt, Efferson, Berger, & Fehr, 2015; Waktare & Roberts, 2014; White, 2015).

As Chapter 1 outlined, there has been a broad consensus since the first eye cues experiments (Burnham, 2003; Burnham & Hare, 2007) that if eye cues work it is by making us feel watched. There is now a significant body of neuropsychological research which suggests this is indeed the case. In their unified theory aiming to explain the mechanism through which eye cues influence behaviour, Conty, George and Hietanen (2016) suggest that watching eyes prompt a subconscious reputation management mechanism to change our behaviour. First, eye cues' highly salient nature captures attention which, second, triggers heightened self-referential processing. In triggering self-referential processing, eye cues are unconsciously prompting heightened concern over how we are socially evaluated. These subconscious reputational concerns can lead us to moderate our behaviour (Conty et al., 2016). In short, this false cue can alter behaviour by changing the way we process information, leading us to act as if are watched and our reputation is at stake even when in fact it is not. If this is so, we should consider the reputational consequences of behaviours when predicting whether eye cues are likely to elicit a strong effect on decision making and also in predicting the direction of any effect.

Having thoroughly reviewed the literature this Chapter argues that the inconsistent results in watching eyes effects studies and in recent meta-analyses may be due to the varying effects of prosocial behaviour on reputation. In contrast this Chapter both proposes and tests the hypothesis that eye cues should more reliably reduce antisocial behaviour, since antisocial behaviour is more consistently reputationally damaging.

This chapter explores the hypothesis that eye cues may consistently reduce antisocial behaviour in four parts. First, it reviews the evidence for the ‘watching eye’ effect on prosocial behaviour highlighting its inconsistency. Second, it provides evidence for the argument that antisocial behaviour may be more consistent in damaging reputation than prosocial behaviour is in enhancing it. Third, it reports the meta-analytic procedure undertaken to test the hypothesis that eye cues will have a more reliable effect on anti-social behaviour before fourth, reviewing the results.

Early reviews in the field suggested there was a reliable influence of eye cues on prosocial behaviour. (Nettle et al., 2013; Sparks & Barclay, 2013). However, Nettle et al.’s (2013) conclusion that in dictator games the presence of eye cues increased not the mean donation but the probability of donating something rather than nothing was recently challenged by Northover and colleagues (2017a) whose larger meta-analysis and review found no effect of eye cues on ‘the proportion who gave’ in generosity tasks across 27 experiments. Furthermore, in a separate analysis Northover and colleagues also found that the effect size was extremely small when comparing the difference between mean donations in the eyes and no eyes conditions across 26 dictator game and charitable giving experiments (Northover et al, 2017a). Northover and colleagues have also investigated the effect of eye cues of moral judgement. They analysed six studies and found no statistically significant eye cue effect, an effect size close to zero and a confidence interval crossing zero (Northover, Pedersen, Cohen, & Andrews, 2017b). These inconsistent findings might also be explained by the variable effects of generosity on reputation.

### 3.1.1. Eye Cues and Antisocial Behaviour

Evidence for the inconsistent effect of generosity on reputation can be found in a series of 11 experiments conducted by Klein and Epley (2014). These show that generosity is not always good for one's reputation. In fact, while some generosity enhances reputation, greater generosity delivers diminishing reputational returns and extreme generosity might even be damaging.

Klein and Epley ran scenarios in which participants rated the reputation of a fictional character "Tom" after reading scenarios that described how much he had donated for a charity concert ticket: below the suggested amount, at the suggested amount or above it. No significant difference was found in ratings of Tom's competence across the 3 levels of donation. This is as expected given the description was identical barring the figure for how much he donated. In contrast, the amount he donated did matter to his reputation for warmth. Behaving fairly by donating as suggested rather than unfairly, donating below the recommended amount, improved ratings of Tom's 'warmth'. But greater generosity did not bring higher 'warmth' ratings. Warmth is related to judgements of friendliness, trustworthiness, and morality and is more dominant in the formation of reputation than competence (Willis and Todorov, 2006; Wojciszke & Abele, 2008).

Klein and Epley conducted additional experiments to further test the relationship between prosociality and reputation via the online crowd-sourcing site Amazon Mechanical Turk, and in laboratory-based settings with similar results. Competence ratings varied little, but warmth ratings did, with extreme generosity decreasing rather than increasing reputational ratings. As the authors put it, ‘it pays to be nice, but pays no more to be really nice’. An extension of this research across 7 countries and cultures has found similar results (Klein, Grossmann, Uskul, Kraus, & Epley, 2015). There is no consistent effect of generosity on reputation, so we should not expect eye cues to cause a consistent increase in generous behaviour.

Surveillance cues, if they provoke reputational concerns, should have more consistent effects in reducing antisocial and selfish behaviour than in promoting prosocial acts because of our acute sensitivity to negative information (Rozin & Royzman, 2001). In an extensive review, instructively titled *Bad is Stronger than Good,* Baumeister, Bratslavsky, Finkenauer, and Vohs (2001) describe a number of experimental findings that suggest why this might be so. They show how we tend to remember bad events longer – thus antisocial acts will echo longer in the memory of the observer than prosocial behaviours. They demonstrate how negative experiences and negative language in relationships exert more influence on ratings of relationship happiness than positive experiences and positive language – thus we might be keener to avoid providing negative signals about ourselves for others to gossip about than we are to seek to provide positive signals and be talked about positively. They demonstrate that negative affect (emotion) and emotional distress have a greater and more enduring influence on us than positive affect and pleasant emotions, so we would be more motivated to avoid the censure that follows signalling bad character than to signal good character to experience warm emotions. They show how we learn faster from negative experiences than positive and how negative events in childhood have greater power than positive events in predicting success in adulthood. They demonstrate how we prioritise the processing of negative over positive information, allocating greater attention and more cognitive resource to it, and how negative stereotypes are more prevalent, more enduring and more influential than positive. They demonstrate how our self-esteem is more sensitive to criticism than it is to praise and how we respond more to critical feedback than to praise. Thus, the evidence from the psychological research summarised by Baumeister et al. (2001) suggests negative information should have consistently damaging – and lasting – effects on our reputation.

The contention that the reputational threat from displaying antisocial behaviour should be greater and more consistent than that from failing to show generous behaviour is supported by evidence from evolutionary psychology too. Punishment in social groups is necessary for group success helping prevent free-riding and social infractions, enforcing hierarchy to enable group functioning and enabling group cohesion (Fehr & Gachter, 2002; Gürerk, Irlenbusch, & Rockenbach, 2006). Indeed, as we have seen an acute fear of being watched may have been essential for social groups to succeed in order to enforce norms without constant combat and punishment. The difficulty of extending surveillance beyond the small groups of our early evolutionary history may have been responsible for the emergence of belief in all-seeing, punishing supernatural agents – gods and spirits – who could ensure we all felt watched even in the absence of others (Johnson, 2016). We should therefore expect stronger, more consistent responses to reputational threat (being seen to behave antisocially) than ambiguous reputational opportunity (being seen to be generous).

In the evolutionary context it is interesting to note that participants in Klein and Epley's experiments expected to earn a better reputation by being more generous. This common-sense assumption underpins experiments examining the effect of watching eyes on generosity. While the assumption that greater generosity generates consistent and linear reputational benefits is unsound it may still be adaptive. Expecting generosity of others creates an explicit social norm that would be of benefit to the group – we should all give, think highly of those who are generous and thus both individual and the group gain if we give when watched. However, an unconscious mechanism promoting prosociality even when not watched would encourage maladaptive behaviours for the individual – leading them to give away resources without prospect of reward.

Given this logic it is unsurprising that cues to surveillance such as watching eyes produce varied responses on people's generosity. In contrast, having an unconscious mechanism to avoid unfairness and antisociality when watched would protect the individual's reputation and thus avoid damaging reprisals. Thus there are sound theoretical reasons to focus our meta-analysis on the effect of watching eyes on anti-social behaviour.

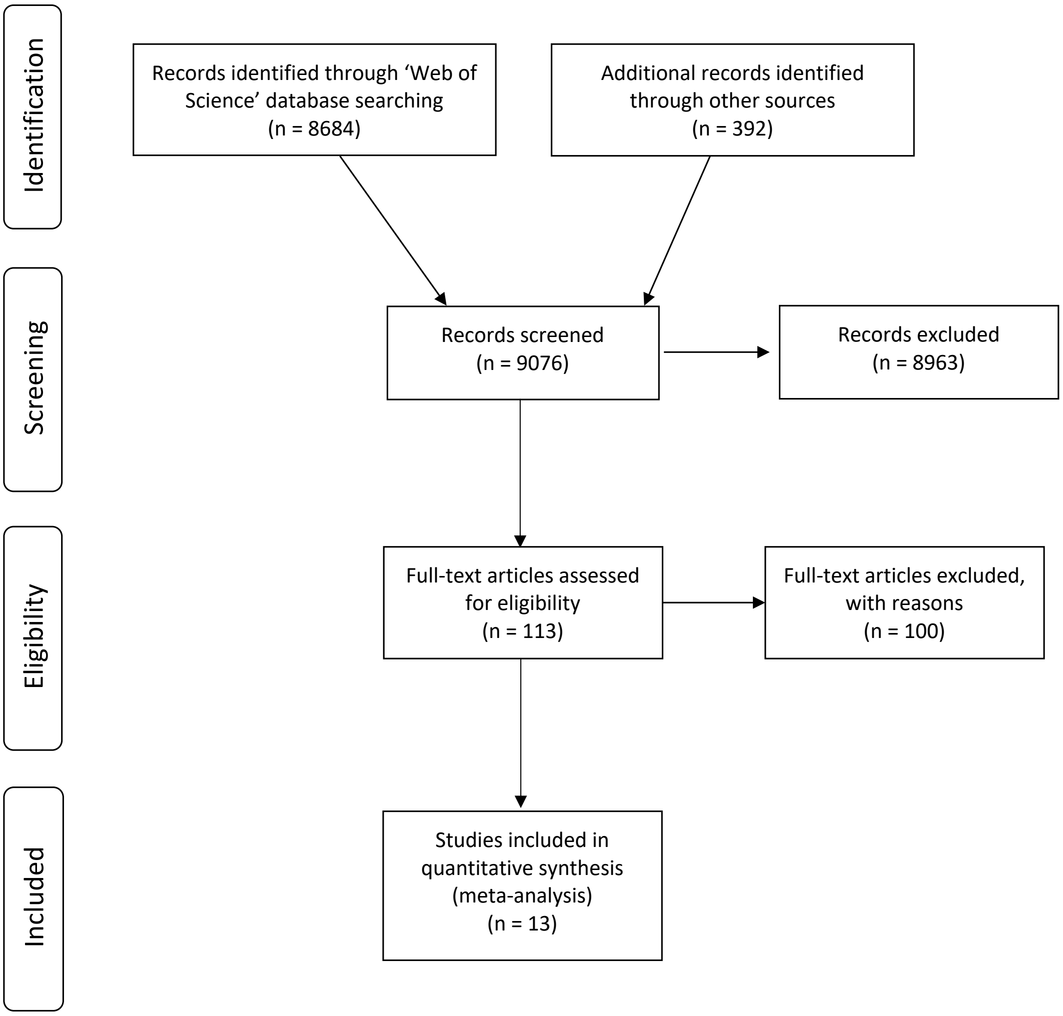
There are sound practical reasons for this focus too. The eye cue effect on antisocial behaviour is already being put to practical use across Britain. In 2006 West Midlands Police used 100 buses to promote the message 'We're Keeping an Eye on Crime' along with a picture of eyes to try to reduce bicycle and other theft (West Midlands Police, 2006). British Police in the Nottinghamshire town of Hucknall claim posters of eyes reduced crime by 40% (Flanagan, 2013). Eye cues have been added to trees to reduce littering in the Forest of Dean, and are being used by Britain's HM Revenue and Customs Service to discourage tax evasion (BBC News, 2015; Knapton, 2016; Nelson, 2013). Press reports indicate that the posters have been used across Britain's rail network to deter crime (BBC News, 2013; Basildon, Canvey & Southend Echo 2013). At the time of writing eye cues are appearing at motorway service stations across Britain as Keep Britain Tidy expands an initiative to reduce littering by drivers, with preliminary results suggesting that they have caused a 23% reduction in littering (Extra Services, 2015; H.M. Government (UK), 2017).

In fact, so well established is the idea in policy circles that the UK Government's 2017 National Anti-Littering Strategy recommends the use of watching eyes interventions to reduce littering (H.M. Government (UK), 2017). Thus, given the strength of the theoretical basis for our hypothesis and its practical implications, we focus our meta-analysis on the watching eyes' effect on antisocial behaviour.

## 3.2. Methods

The meta-analytic model followed the PRISMA flow model as per Figure 3.1.

*Figure 3.1. PRISMA Flow Diagram after Moher, Liberati, Tetzlaf, Altman and Group (2009) and Shamseer et al. (2015).*


Identification
Screening
Eligibility
Included
Records identified through ‘Web of Science’ database searching (n = 8684)
Additional records identified through other sources (n = 392)
Records screened (n = 9076)
Records excluded (n = 8963)
Full-text articles assessed for eligibility (n = 113)
Full-text articles excluded, with reasons (n = 100)
Studies included in quantitative synthesis (meta-analysis) (n = 13)

### 3.2.1. Types of intervention

We reviewed 113 full-text articles comparing between groups exposed to eye cues and those in a control condition. We examined both laboratory and field experiments. We considered for meta-analysis 19 articles identified as examining the watching eyes effect on antisocial behaviour, excluding a further 6 due to issues with the experimental design or the unavailability of data as described below.

### 3.2.2. Types of outcome measures

The outcome measure for the meta-analysis was antisocial behaviour.

### 3.2.3. Search strategy

Initially, we conducted a full search without limitation by outcome measure to document all eye-cue relevant research in one place for the first time. In order to take as full a view of the watching eyes field as possible we searched Web of Science by topic using the search terms "watching eyes", "eye cues", "eye cue", "eye spot", "eyespot", "eye-like", "social cues", "eye-images", "watching you", "surveillance cues", "images of eyes", "observation cues", "perception of human face", "face cues", "cues of observation", "gaze cues", "implicit gaze", "implicit observability", "eye primes", and by title using the terms "implicit social", "watching you", "social eyes OR social-eyes" (See Additional Materials for full details of our search strings). We excluded duplicates counting only studies not previously identified using earlier search terms. We identified studies in the bibliography of the most recent and thorough review (Northover et al., 2017a).

Noting that Northover et al. had just completed an extensive search for unpublished research at the same time as our review was written we did not duplicate this effort by announcing our own formal call for unpublished papers. However, we did conduct our own private search for unpublished material - searching by author name when alerted to their interest in the watching eyes effect in other papers - and in doing so identified 4 additional studies to those picked up by keyword searches. Two further studies (Palomäki et al., 2015; Traver & Cordell, 2014) were identified via the abstract booklet of the *Decepticon* conference attended by a colleague. Palomäki et al.’s unpublished study was received via correspondence with the author and is included in the meta-analysis.

We searched studies using Google Scholar’s ‘cited by’ function to identify further research in different domains that we might not otherwise have found (generosity in field and lab, voting behaviour, minimal cues), identifying a further 15 studies. Seeking real-world applications of the research we conducted a general web search using Google, finding a further large-scale field experiment (and a number currently underway).

An extensive search was made of Newcastle University’s online repository (http://research.ncl.ac.uk) which yielded a number of unpublished Masters’ research studies.

We repeated our search prior to submitting for publication in March 2017 adding a further 13 studies.

All searches are documented in the Excel tabulation in ‘Additional Materials’ and provide a comprehensive overview of the 48 different outcome measures identified ranging from the watching eyes effect on generosity and hand-washing to their influence on choice justification and poker risk-taking.

This is, to the best our knowledge, the most comprehensive database of watching eyes studies yet compiled.

### 3.2.4. Meta-Analysis Study Selection

All studies containing the search terms were reviewed by title and abstract. Where their relevance remained unclear the full text was accessed. Studies were eliminated from, or included in the review in this way on the first search. Further refinement for the meta-analysis was undertaken following the recording and categorisation of the studies [See Excel spreadsheet in additional materials].

With all watching eyes studies documented in Excel format we filtered the table to identify watching eyes studies examining antisocial behaviour, selecting from the column 'outcome measure' the antisocial behaviour categories: 'Advice Giving', 'Amount Taken', 'Antisocial Behaviour', 'Cheating', 'Corruption', 'Cycle Theft', 'Dishonesty', 'Dog fouling', 'Engine Idling', 'Honesty', 'Littering', and 'Lying'. This selection process returned 19 papers. All others were excluded as they did not examine antisocial behaviour [this process can be replicated via the Excel database in additional materials]. The 19 papers selected all studied behaviours that met the Oxford English Dictionary definition of antisocial, that is, 'contrary to the laws and customs of society, in a way that causes annoyance and disapproval in others'.

From these 19 papers we excluded six. Bateson, Nettle, and Roberts' 2006 study of the effect of images of watching eyes on donations to a coffee bar honesty box was excluded because its design did not allow us to know whether the experiment examined conformity to norms, antisocial behaviour or generosity. It did not, for example, count or track how much individuals donated, or how many did or did not donate. We therefore cannot know whether eye cues increased the generosity of those who would have donated anyway or whether eye cues caused people who would have free-loaded and failed to donate, to contribute. We are also unable to ascertain whether this was a test of the eye cue effect on antisocial behaviour or whether their result was a combination of both the eye cue effect increasing generosity in some and reducing antisocial free-riding in others. The same criteria resulted in the exclusion of Bruderman et al.'s (2015) field experiment examining contributions to newspaper honesty boxes.

We excluded Oda, Kato, and Hiraishi's (2015) study examining the eyes effect on lying principally because their design construed lying as pro-, as opposed to antisocial behaviour.

Traver and Cordell's paper (2014) was excluded as we could not create a dichotomous variable suitable for our odds ratio-based meta-analysis. These authors examined the eye cue effect in a coin toss and a dice roll experiment (where flipping a head or rolling a higher number brought a greater reward) reporting the number of outcomes against the expected frequencies to identify how frequently people lied to gain personal advantage. We could not calculate the frequency of lying vs not lying from the proportions as all frequencies lay within one standard deviation of the expected frequencies, and thus could have been down to chance.

One study (Chowdhury, Jeon, & Saha, 2014), was excluded as the authors were unwilling to share the data on which their study was based because they were seeking re-publication of their paper in another journal. There was insufficient information reported in the available paper to assess dichotomously who did and did not behave antisocially [data requested on 12 Oct 2016 and refused in writing on 13 October 2016 & again on 23 August 2017]. Data was requested from the authors for Li, Zhan, Fan, Liu, Li, Sun, Zhong, (2018) but no reply was received at the time of submission.

The 13 studies remaining reported results for a total of 15 eligible experiments. We transformed all findings into dichotomous variables to test the simple effect of eye cues on antisocial behaviour via the log odds ratio. This measure is intuitively meaningful because it indicates the relative change in frequency of antisocial behaviour in the experimental condition or area compared with the control condition or area. For example, in the study of littering in a café by Ernest-Jones, Nettle, and Bateson (2011) we transformed the data into dichotomous 'littered or did not litter' variables from the original data. This comes at the cost of hiding significant interactions e.g. in the Ernest-Jones' et al. example they found a significant effect of eye cues but only when >6 people were present. Such an effect is hidden when we transform the data to the dichotomous did or did not litter by eyes and control.

Within some of the papers selected some experiments were eligible for inclusion in the meta-analysis while others were not. For example, in Zengerink (2013) Experiment 1 involved the completion of a series of surveys with eye cues present but did not examine antisocial behaviour and was therefore excluded. Experiment 2 was a field experiment examining the effect of eye cues on antisocial behaviour and therefore was included. The same approach was taken with two other experiments examining the eye cue effect on litter (Bateson et al., 2015; Zengerink, 2013).

We reinterpreted the data from Cai et al. (2015) examining whether eye cues affected participants' tendency to lie for personal advantage to show the frequency of lying or not lying in the eyes and control condition. Palomäki et al. (2015) also examined the effect of eye cues on lying, in their case in an online simulated insurance claims process. We re-analysed their data to provide a dichotomous outcome: lied or did not lie. We reinterpreted Hoffman et al.'s data from their study of the effect of robot eyes on lying (2015) to provide dichotomous proportions of those that cheated or did not cheat and reinterpreted the proportions as frequencies.

Huang et al. (2015) examined whether participants would bribe or not bribe in scenario-based economic games studying corruption. Their Experiments 2 and 3 include an 'eye cue or no eyes' condition enabling this data to be translated for meta-analysis.

In Baillon et al.'s 'Joy of Destruction' (2013) game participants could pay a small cost to destroy part of a payoff to another participant gaining nothing in the process except the possible satisfaction of damaging the other's interest. We transformed their data to report a dichotomous outcome "destroyed" or "did not destroy" in the eyes and control conditions.

Finally, in an extended analysis, we included Nettle, Nott and Bateson's (2012) examination of the effect of eyes on bicycle crime and Keep Britain tidy's (2014) eye cue intervention designed to reduce dog fouling. In Nettle et al. (2012), Bateson et al. (2013), and some of Keep Britain Tidy's (2014) crime reduction interventions the eye images were used along with text explicitly stating that participants were being watched. For example, in Nettle et al.'s study the eye cues sign included the text 'Cycle Thieves We're Watching You' and a police logo. Bateson et al. (2013) used the similar signs as Nettle et al. (2012) in their experiment. Similarly, Keep Britain Tidy's intervention posters all included text and various organisational logos. Consequently, it might be argued that either the logos or the text alone, or the two in combination, caused the reduction in the bicycle theft, and not the watching eyes. Indeed just erecting signs alone can reduce crime (McNees, Egli, Marshall, Schnell and Risley, 1976) and so we may not be able to attribute any crime reduction to the presence of eye cues in these two studies. Additionally, it is plausible that people seeing the signs could have thought they suggested CCTV was present, which perhaps might also reduce crime although the evidence for the crime-reducing effects of CCTV is inconsistent (Welsh, 2009).

One could make a case for excluding Nettle et al. (2012), and some of Keep Britain Tidy's (2014) studies due to the presence of these potential confounds. However, given that Nettle et al. (2012) is perhaps the most influential paper in this area on public policy in the UK, and also represents one of the baseline studies in the research field, many would be surprised to see it excluded. Indeed, reviewers have offered conflicting advice on this point.

Consequently, we calculate log odds ratios with and without the inclusion of Nettle et al. (2012), Bateson et al. (2013) and Keep Britain Tidy (2014) to satisfy both perspectives.

Keep Britain Tidy's study had 4 types of poster, all with eyes, but with differing text to examine whether this increased or decreased the effectiveness of the intervention. As well as having pictures of eyes, these four poster designs also displayed the text 'Thoughtless Dog Owners, We're Watching You'. While the first type had no further text, the three others added different messages to try to select the most effective wording to deter dog owners from allowing their dogs to defecate in the area. Because in this meta-analysis we are interested only in the main effect of eyes, we collapsed the data for the four eyes posters into one experimental condition.

Furthermore, we reinterpreted Keep Britain Tidy's 'displacement areas' as controls as they were functionally equivalent to Nettle et al.'s (2012) control areas.

### 3.2.5. Analytic Methodology & Statistical Approach

For the first of our two analyses we identified studies comparing the number of participants behaving antisocially or not in an eyes and a control condition in order to enable the calculation of the log odds ratio as a measure of effect size. The odds ratio is used in one of the landmark 'watching eyes and antisocial behaviour' experiments by Nettle et al. (2012) and was selected as the measure of effect size as recommended by Welsh and Farrington (2009) for meta-analysis of location-based crime interventions, and by Lipsey and Wilson (2001) for meta-analysis of inherently dichotomous variables. Our statistical methods used in the fixed effects and random effects analyses mirror those of Northover et al. in their two meta-analyses of the watching eyes effect on generosity (2017a) and on moral judgement (2017b), thus enabling comparison.

Odds ratios are centred around 1 rather than zero (where 1 indicates no relationship, values 0-1 a negative relationship and values greater than 1 a positive relationship). The log odds ratio is distributed approximately normally with a mean of zero and a standard deviation of 1.83, thereby allowing us to represent a negative effect (a reduction in antisocial behaviour when eyes are present) with a negative value and a positive effect with a positive value, thus simplifying interpretation and calculation.

| | *Antisocial Behaviour (ASB)_YES* | *ASB_NO* |
|---|---|---|
| Eyes Condition (Group 1) | (a) | (b) |
| Control (Group 2) | (c) | (d) |

The log odds ratio effect size is calculated after Lipsey and Wilson (2001):

$$OR = \frac{(a * d)}{(b * c)}$$

*Equation 1*

$$OR_{log} = log_e(OR)$$

*Equation 2*

The logs odds ratio standard error and inverse variance weight are calculated as:

$$SE_{LOR} = \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}$$

*Equation 3*

$$W_{LOR} = \frac{1}{SE_{LOR}^{2}}$$

*Equation 4*

We applied Cochran’s *Q* test to estimate whether the individual effect sizes were representative of the population values:

$$Q = (\sum W_i \, ES_i^2) - \frac{(\sum W_i \, ES_i)^2}{\sum W_i}$$

*Equation 5*

Finding a statistically significant *Q* indicates a heterogeneous distribution greater than that expected from subject level sampling error alone. A significant *Q* value therefore suggests it may be appropriate to apply a random effects model to account for variance introduced by study differences such as varying experimental designs or outcome measures. We first calculated the random effects variance ($V_\theta$).

$$V_\theta = \frac{Q - (k - 1)}{\sum W_i - (\sum W_i^2 / \sum W_i)}$$

*Equation 6*

Where $k$ is the number of effect sizes, and $W_i$ is the inverse variance weight for each effect size. The random variance was then added to the sampling error and the inverse variants weights recalculated as:

$$\frac{1}{(V_\theta + SE_{LOR}^2)}$$

*Equation 7*

The log odds ratio was then recalculated using the new weights.

Next, we calculated the lower and upper bounds of the 95% confidence interval of the log odds ratio:

$$95\%\ CI\ Lower = OR_{log} - (1.96 * SE_{LOR})$$

*Equation 8*

$$95\%\ CI\ Upper = OR_{log} + (1.96 * SE_{LOR})$$

*Equation 9*

We calculated the *z*-statistic and *p*-values for our log odds ratios:

$$Z = \frac{OR_{log}}{SE_{LOR}}$$

*Equation 10*

$P$-values were than calculated after the method recommended by Altman and Bland (2011):

$$P = \exp(-0.717 * Z - 0.416 * Z^2)$$

*Equation 11*

However, Nettle, Nott and Bateson's (2012) field study examining the eyes effect on bicycle theft could not report how frequently people decided *not* to behave antisocially (i.e. how many were 'exposed' to the experimental and control conditions and chose *not* to steal bikes). Instead, they report only the frequency of antisocial behaviour across experimental and control locations. On account of this we calculated both a log odds effect size and, additionally selected and report the incidence ratio (sometimes called the rate ratio) as an alternate measure of effect size. Such a rate ratio is calculated by controlling for any change in behaviour in both experimental and control locations before and after the intervention (CTSPedia, 2009).

Similarly, the experiment run by Keep Britain Tidy (2014) examined only before and after counts of dog faeces found at each of the 240 sites where signs were put up, comparing these the total number of dog faeces found over the same before and after period at the 240 adjacent control areas (or displacement areas, as they termed them)

| | Before (ASB = Bicycles stolen or dog faeces in area) | After (ASB = Bicycles stolen or dog faeces in area) |
|---|---|---|
| Experimental Locations (sheds or areas where eyes signs were introduced for the second 3 weeks of the six week experiment) | (a) | (b) |
| Control Locations (sheds or areas where no signs were introduced during the six weeks of the experiment) | (c) | (d) |

$$Incidence\ Rate\ Ratio = \frac{a/b}{c/d}$$

*Equation 13*

Analyses were performed in *R* 3.1 using the METAFOR package’s random effects model function. *R* Script is provided in Additional Material to aid with replication of our analysis (*R* Development Core Team, 2016). Calculation of 95% confidence intervals, *Z*-scores and *P*-values was completed in Microsoft Excel. *P*-curve analysis was conducted after Simonsohn, Nelson, and Simmons (2014) and (2014b) using their online app provided at http://www.p-curve.com. PET & PEESE analysis was undertaken in R using code provided by Hilgard (2016).

## 3.3. Results

Table 3.1. summarises out metanalytic output, providing log odds ratios and associated statistics for the dichotomous data from all studies included in the analyses that follow.

**Table 3.1.** ***Studies in the Meta-Analysis Exploring the Effect of Eye Cues on Antisocial Behaviour (ASB)***

| | | | **Eye Cues Present** | | | **Control** | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Citation & Experiment Number** | **Outcome Variable** | ***N*** | **ASB Yes** | **ASB No** | ***N*** | **ASB Yes** | **ASB No** | $OR_{log}$ | $Var_{LOR}$ | $SE_{LOR}$ | **95% CI Lower** | **95% CI Upper** | ***Z*** | ***P*** |
| **1** | Ernest-Jones et al. (2011) | Litter | 273 | 56 | 233 | 289 | 98 | 175 | -0.85 | 0.04 | 0.20 | -1.23 | -0.46 | -4.34 | **0.01** |
| **2** | Bateson et al. (2013) | Litter | 305 | 80 | 225 | 297 | 83 | 214 | -0.09 | 0.03 | 0.18 | -0.45 | 0.27 | -0.47 | 1.28 |
| **3** | Baillon et al. (2013), Experiment 1 | Destruction of others money | 51 | 9 | 42 | 49 | 19 | 30 | -1.08 | 0.22 | 0.47 | -2.00 | -0.16 | -2.31 | 0.57 |
| **4** | Zengerink (2013), Experiment 2 | Litter | 630 | 297 | 333 | 314 | 170 | 144 | -0.28 | 0.02 | 0.14 | -0.55 | -0.01 | -2.02 | 0.78 |
| **5** | Bateson et al. (2015), Experiment 1 | Litter | 147 | 8 | 139 | 137 | 23 | 114 | -1.25 | 0.18 | 0.43 | -2.10 | -0.41 | -2.92 | 0.23 |
| **6** | Bateson et al. (2015), Experiment 2 | Litter | 216 | 31 | 185 | 97 | 25 | 72 | -0.73 | 0.09 | 0.30 | -1.32 | -0.14 | -2.41 | 0.50 |
| **7** | Cai et al. (2015), Experiment 1 | Dishonesty | 66 | 30 | 36 | 65 | 31 | 34 | -0.09 | 0.12 | 0.35 | -0.78 | 0.60 | -0.26 | 1.17 |
| **8** | Cai et al. (2015), Experiment 3 | Dishonesty | 64 | 18 | 46 | 66 | 22 | 44 | -0.25 | 0.15 | 0.38 | -0.99 | 0.50 | -0.64 | 1.34 |
| **9** | Hoffman et al. (2015) | Dishonesty | 20 | 7 | 13 | 20 | 9 | 11 | -0.42 | 0.42 | 0.65 | -1.69 | 0.85 | -0.64 | 1.34 |
| **10** | Palomäki et al. (2015) | Dishonesty | 105 | 13 | 92 | 86 | 9 | 77 | 0.19 | 0.21 | 0.46 | -0.71 | 1.09 | 0.41 | 0.69 |
| **11** | Huang et al. (2015), Experiment 2 | Corruption | 24 | 10 | 14 | 24 | 9 | 15 | 0.17 | 0.35 | 0.59 | -0.98 | 1.33 | 0.30 | 0.78 |
| **12** | Huang et al. (2015), Experiment 3 | Corruption | 30 | 11 | 19 | 31 | 24 | 7 | -1.78 | 0.33 | 0.57 | -2.90 | -0.66 | -3.11 | 0.17 |
| **13** | Meleady et al. (2017), Experiment 1 | Engine idling | 112 | 83 | 21 | 104 | 82 | 39 | 0.63 | 0.10 | 0.31 | 0.02 | 1.24 | 2.02 | **0.04** |
| | | **Experimental Locations** | | | | | **Control Locations** | | | | | | | | |
| | **Citation & Experiment Number** | **Outcome Variable** | ***N*** | **Frequency of ASB Before** | **Frequency of ASB After** | ***N*** | **Frequency of ASB Before** | **Frequency of ASB After** | $OR_{log}$ | $Var_{LOR}$ | $SE_{LOR}$ | **95% CI Lower** | **95% CI Upper** | ***Z*** | ***P*** |
| **14** | Nettle et al.(2012) | Bicycle Theft | N/A | 31 | 51 | N/A | 39 | 15 | -1.45 | 0.14 | 0.38 | -2.20 | -0.71 | -3.83 | **0.04** |
| **15** | Keep Britain Tidy (2014) | Dog Fouling | N/A | 861 | 434 | N/A | 2159 | 1208 | 0.10 | 0.00 | 0.07 | -0.03 | 0.24 | 1.51 | 0.13 |

We conducted a homogeneity analysis test to examine the assumption that all of our effect sizes were estimating the same population mean. The fixed effects model on the data without Nettle et al. (2012), Bateson et al. (2013) and Keep Britain Tidy (2014) returned a result of $OR_{log} = -.42$, (SE = .08, 95% CI [-.59 to -.26]); exponentiating the $OR_{log}$ we get an OR = 0.65, suggesting a 35% reduction in antisocial behaviour by condition. However, since our $Q$-test for homogeneity of variance was significant ($Q(11) = 33.99$, $p < .001$), suggesting significant heterogeneity across our studies it proved necessary to fit a random effects model to the data.

The random effects model produced a meta-analytic effect size of $OR_{log} = -.45$, (SE = .18, 95% CI [-.81, -.09]) $p = .01$. Transforming our effect size to aid with interpretation we found an odds ratio (via $OR = e^{OR_{log}}$) of .64, suggesting a 36% reduction in the risk of antisocial behaviour in the eyes condition.

Extending our analysis to include Nettle, Nott and Bateson's bicycle crime experiment (2012), Bateson et al.'s littering experiment (2013) and Keep Britain Tidy's (2014) dog fouling intervention, we expected, and found, greater heterogeneity rendering the fixed effects model an inappropriate measure. In the Random Effects Model there was little change in the effect size since Keep Britain Tidy's large-scale experiment found a small increase in antisocial behaviour in the eyes areas in comparison with the control which was sufficient to offset the large effect size in the opposite direction found in Nettle, Nott and Bateson's bicycle crime experiment.

Our fixed effects model including these studies returned a result of $OR_{log} = -.13$, (SE = .05, 95% CI [-.23, -.03]); exponentiating the $OR_{log}$ we get an $OR = .88$, suggesting a c. 12% reduction in ASB in the eyes when compared with the control condition. The $Q$-test for homogeneity of variance was significant ($Q(14) = 68.54$, $p < .01$) again requiring us to fit a random effects model to the data.

The random effects model suggested that there was a 35% reduction in the risk of antisocial behaviour between the eyes and control condition ($OR_{log} = -.43$, ($SE = .17$, 95% *CI* [-.76, -.11]) $p < .01$; $OR = .65$). Neither a forest or funnel plot show any evidence of over-dispersion (See Figures 3.1. and 3.2. below). In addition, a *trim and fill* test for missing values (after Duval and Tweedie, 2000) finds no evidence to suggest either missing reports or publication bias.

The Rate Ratio for all studies including Nettle et al. and Keep Britain Tidy showed a larger effect still, $RR = 0.27$, $SE = 0.09$, $Z = 2.84$. CI [0.08, 0.46], $p < .01$. This suggests a 73% reduction in antisocial behaviour between control and eyes conditions. This seems implausibly large and hence we rely on our more conservative log odds estimates.

*Figure 3.2. Forest plot of experiments included in both the limited and extended meta-analyses. Plot is centered on zero (null effect) and displays effect size in $OR_{log}$ with corresponding 95% CI.*

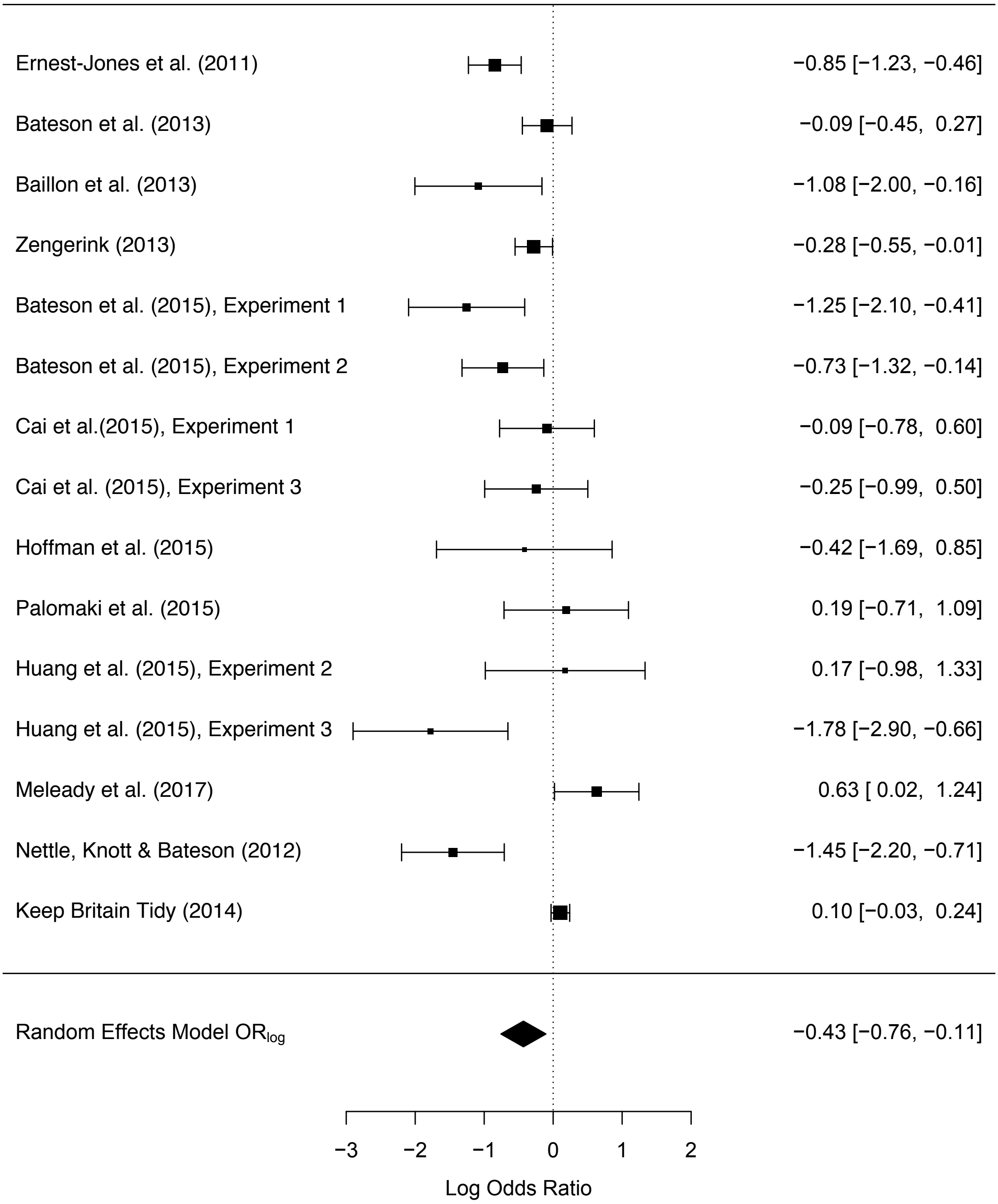
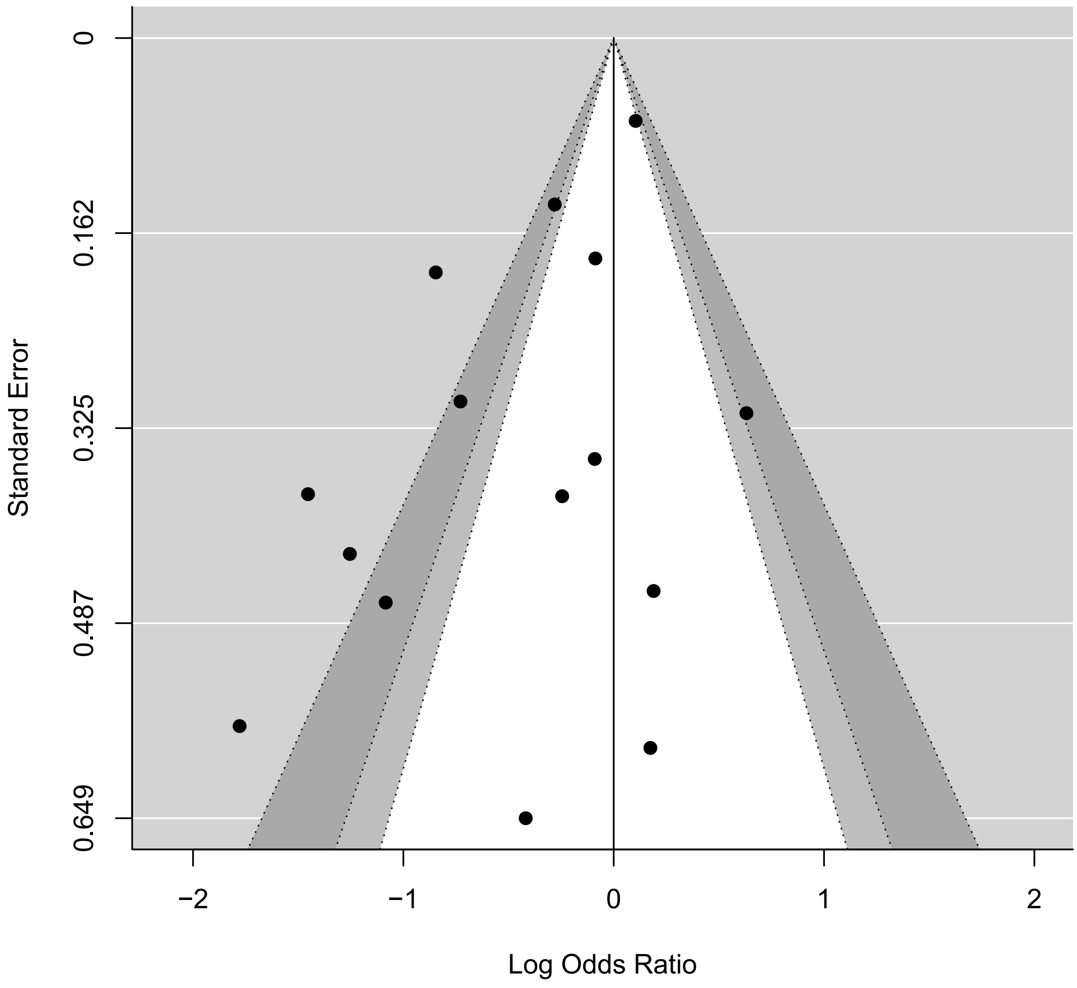


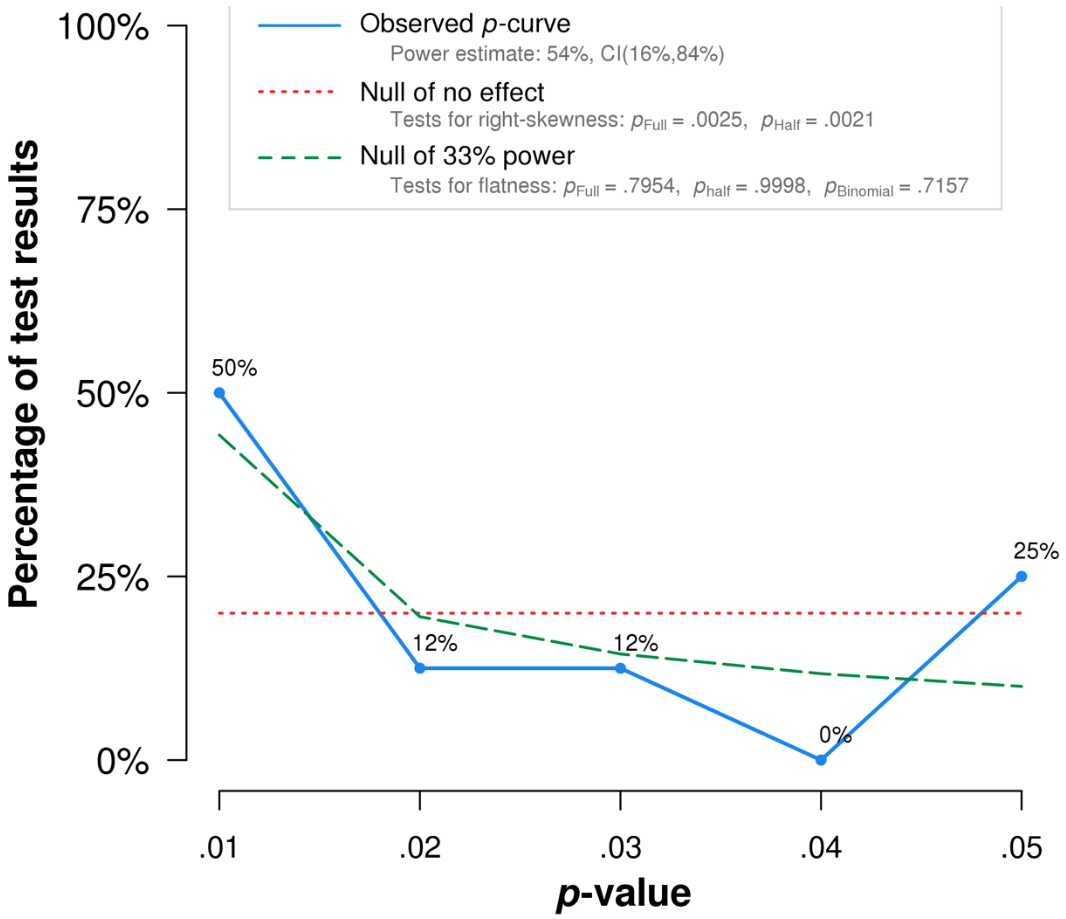
*Figure 3.3. Funnel plot of experiments included in both the limited and extended meta-analyses. Plot is centered on zero (null effect).*



The *Trim and Fill* method (Duval & Tweedie, 2000) relies on the assumption that selective reporting of studies, that is publication bias, is driven by effect size. This is vulnerable to criticism since the principle barrier to publication is widely acknowledged to be achieving statistical significance below $p$ = .05. $P$-curve analysis explores the consistency and distribution of statistically significant results and outperforms other methods in identifying publication bias (Simonsohn, Nelson, and Simmons 2014a & 2014b).

To enable a $p$-curve analysis we calculated $z$-scores from our log odds ratios. Since our analysis relied on data transformed to provide dichotomous results the $p$-values in original published studies were not relevant. We ran this test inputting to the $p$-curve application the *z*-scores for the fifteen studies in Table 1. The application recalculates statistical significance based on the Z-scores alone. Figure 3.3. shows 8 studies of 15 that were statistically significant under this re-analysis. The full output from the *p-curve analysis* can be obtained via Appendix D at the end of this thesis.

*Figure 3.4. P-curve results examining z-scores of our 8 studies for publication bias.*



Note: The observed *p*-curve includes 8 statistically significant ($p < .05$) results, of which 6 are $p < .025$. There were 7 additional results entered but excluded from *p*-curve because they were $p > .05$.

*P*-Curve analysis allows us to draw inferences about evidential value. If the half *p*-curve test – which is much more robust to ambitious *p*-hacking than the simple full *p*-curve test (Simonsohn, Nelson, & Simmons, 2014) - is right-skewed with $p < .05$, or both the half and full test are right-skewed with $p < .1$, then *p* -curve analysis indicates the presence of evidential value (Simonsohn, Simmons, & Nelson, 2015). Here both conditions are met indicating evidential value, as they were when entering only our statistically significant results.

Similarly, *p*-curve analysis indicates that evidential value is inadequate or absent if the 33% power test is $p < .05$ for the full *p*-curve, or both the half *p*-curve and binomial 33% power test are $p < .1$. Here neither condition is met so the *p*-curve does not indicate that evidential value is inadequate nor absent. The same also proved true when entering only our statistically significant studies. This *P*-curve analysis provides some reassurance that our findings are not the product of publication or other bias.

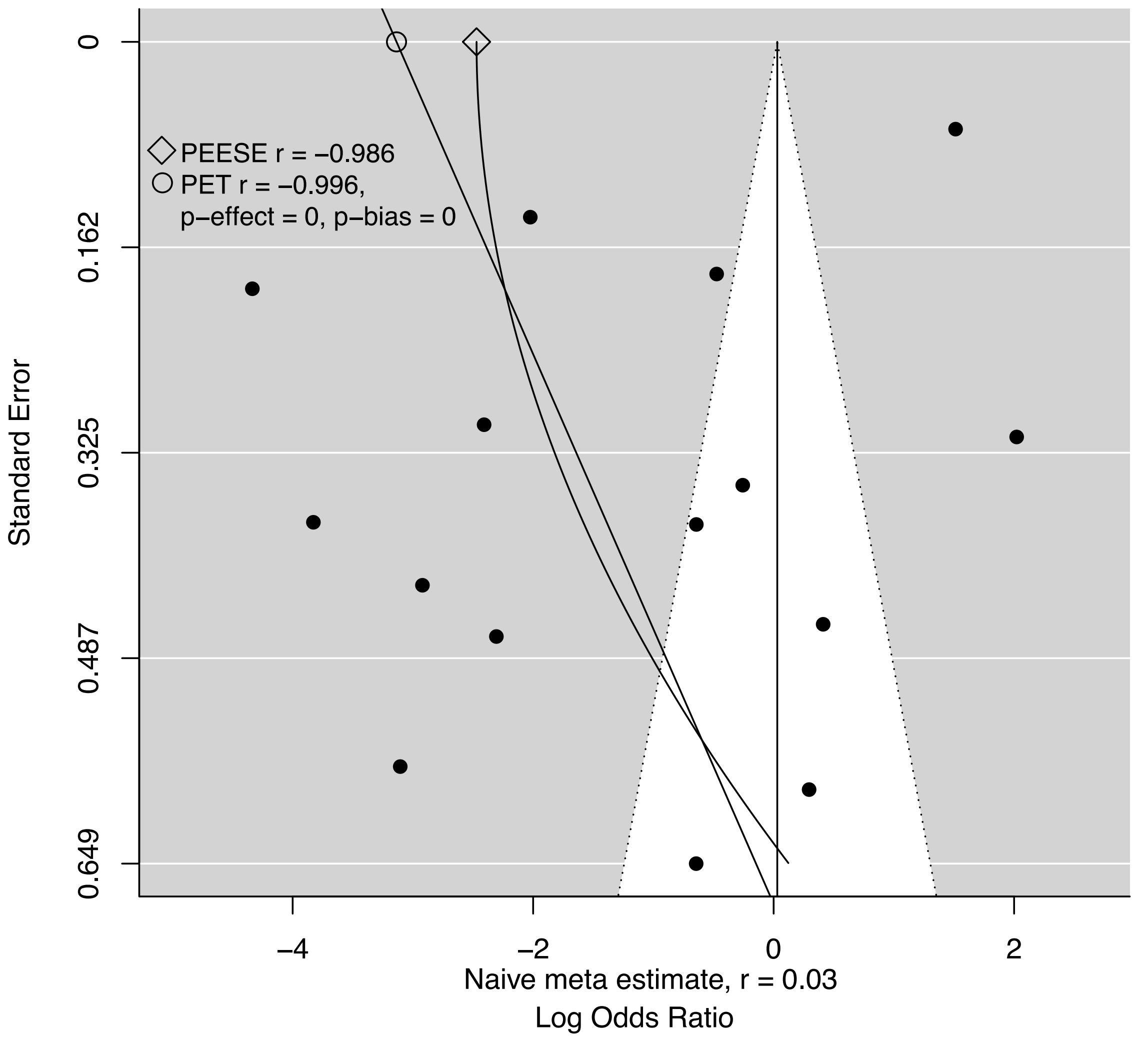
As a further confirmation check on our results we conducted a Precision Effects Test (PET; Stanley 2005) which estimates effect size via a weighted least squares regression model and which provides '*…a valid basis for determining whether there is a genuine empirical effect beyond publication selection bia*s.' (Stanley & Doucouliagos, 2013). Because PET can over-correct for publication bias, underestimating the true effect where one exists (i.e. it is most accurate when the effect size is zero) (Stanley, 2007) it is best treated as a pass or fail test rather than an estimator of effect size (Stanley & Doucouliagos, 2012, pp. 63, 78-79). We therefore added also a Precision Effect Estimate with Standard Errors (PEESE) which uses variance (standard error squared) as a predictor instead of the standard error (Stanley & Doucouliagos, 2007). PEESE is more accurate when there is a true effect (i.e. the true effect is non-zero) and should be selected over the PET output as the more accurate estimate of effect size.

Since random effects models skew PET & PEESE analysis in the direction of any publication bias, we chose a Fixed Effects Model following the advice of Stanley and Doucouliagos (2012, p. 83). Because both PET & PEESE require an *N* value as input we excluded Nettle et al. (2012) and Keep Britain Tidy (2014) from this analysis. We also exclude Bateson et al. (2013) in line with our previous analysis. The *R*-code and our full PET-PEESE analysis are available at Open Science – see note at Appendix D.

Using the Standard Error of the Log Odds ratio and *Z*-Scores (Table 1) in a PET analysis with a Fixed Effects Model and *R* Code from Hilgard (2016), our results suggested that the presence of watching eyes significantly decreased the frequency of antisocial behaviour, $b_0 = -3.14$, $SE = 0.18$, 95% CI[-3.49,-2.78], $p < .0001$, r = -.99.

Our fixed-effects PEESE analysis (after Hilgard, 2016) again found a significant effect, $b_0 = -2.47$, $SE = .12$, 95% CI[-2.70, -2.24], $p < .0001$, $r = -.99$ where $r = .50$ is considered a large effect against conventional heuristics (Cohen, 1992) again suggesting that the presence of watching eyes significantly decreases antisocial behaviour. See Figure 3.4. for a funnel plot of the Fixed Effects PET-PEESE model.

*Figure 3.5. Funnel plot of the Fixed Effects PET-PEESE analysis for experiments 1-13. Plot is centered on the naive meta estimate and displays effect size on the x-axis in* $OR_{log}$



## 3.4. Discussion

In this section we first seek to interpret our findings, second reconcile our meta-analytic output with that of Northover et al. (2017a) and third, extend our eye cues and antisocial behaviour analysis descriptively.

Though our meta-analytic effect sizes are quite widely dispersed they consistently show that the presence of eyes cues correlate with a reduction in antisocial behaviour. Our first fixed effects analysis suggests there is a 35% reduction in the risk of antisocial behaviour in the presence of watching eyes. The Random Effects Model puts the effect at 36%; adding in Nettle et al. (2012) and Keep Britain Tidy (2014) leads the Fixed Effects Model output to suggest the presence of watching eyes may reduce antisocial behaviour by 12%; and the Random Effects Model including the additional studies puts the reduction at 35%. Given the difference in experimental design between Nettle et al. (2012) and Keep Britain Tidy (2014) and the other studies included in our analysis we report our overall effect size as a 35% reduction in the risk of antisocial behaviour in the presence of eye cues based on our first, more carefully selective analysis and using the more restrictive fixed effects model to report the more conservative result.

Our *p*-curve and *trim and fill* analyses find no evidence for publication bias. Our PET-PEESE analysis finds a significant effect in the same direction as our previous analysis – the presence of eye cues correlates with a reduction in antisocial behaviour.

However, there are reasons to be cautious of relying on PET-PEESE. Stanley (2017), the originator of the method, highlights the limitations of the analysis and related methods for small sample studies, areas with few studies, and areas with high heterogeneity. Simulations suggest PET-PEESE misses real effects around 80 per cent of the time, and when heterogenous effects are present but real (i.e. there is not a single underlying effect that never differs regardless of context, and heterogeneity is not just noise and measurement error) or when publication bias is present, PET-PEESE under-estimates the true effect size 60-80% of the time (Gervais, 2016). The analysis also performs poorly when sample size is not reasonably consistent (Hilgard, 2015; Simonsohn, 2017).

The watching eyes literature is characterised by wide variations in sample size, includes many small samples, and finds heterogenous effects that may well be due to context and individual differences not just measurement error and noise. However, as an additional method to check our results rather than the sole analysis, we judge that the results are worthy of inclusion.

We were surprised to find no evidence of publication bias in our statistical analyses. However, perhaps we should not have been. Northover et al. (2017a) had conducted a search for missing papers concurrent within the period that we completed our meta-analysis finding just 2 unpublished studies. Our own extensive search for unpublished studies found only four additional papers. It may be that there are many more watching eyes studies in the file drawer, but neither extensive web-searching, specific calls for unpublished papers nor statistical analysis has yet been able to find any evidence for this.

Northover et al.’s meta-analysis (2017a) is the benchmark for meta-analysis in the watching eyes field. It finds null effects where our effects are significant. Consequently, we shall orientate our discussion around it. However, it is important to note that, valuable as Northover et al.’s analysis is there are two weaknesses within it. Firstly, their decision to include two studies from Tane and Takezawa (2011) which were conducted in darkness. These two studies were implicitly testing the boundary conditions of the watching eyes effect rather than its existence – if watching eyes cue people to feel watched presumably darkness makes people feel less conspicuous. As noted in Chapter 1, there is evidence to support this common-sense assumption. Zhong, Bohns & Gino (2010) report that darkness can induce illusory anonymity leading to more cheating in economic games. Therefore, the inclusion of Tane & Takewaza’s studies skews the analysis. Secondly, Mifune et al.’s (2010) experiment is included in Northover’s analysis, but this study was explicitly conducted to show the boundary conditions where eye cues would *not* work. The outcome of the experiment was as predicted by Mifune and colleagues: eye cues increased generosity to in-groups but not to out-groups. However, this nuance is lost by simply including their study as ‘null results’.

Finally, it should also be noted that no-one, including Northover and colleagues, suggests that their meta-analysis provides definitive evidence that eye cues have no effect on generosity.

Our meta-analysis finds a large effect size, indicating a reduction in antisocial behaviour in the presence of eye cues. In marked contrast, in both of the meta-analyses by (Northover et al., 2017) a vanishingly small effect of eye cues on generosity was found. This is true both when comparing the standardised mean difference and when examining the probability of donating something rather than nothing.

How, then, do we reconcile our findings that eye cues have a large effect in reducing antisocial behaviour with those of Northover et al. that show no effect of eye cues on generosity?

In seeking to reconcile our findings with Northover et al.'s, we first consider recent experimental evidence that upholds the intuitive idea that eye cues cause people to feel watched. If, as the evolutionary legacy hypothesis suggests, eye cues make us feel watched, and thereby make us more aware of our reputation, then the difference between Northover's meta-analytic findings and our own may be best explained by the variable effect of generosity on reputation and the more consistent effect of antisocial behaviour on reputation, as outlined earlier.

Pfattheicher and Keller (2015) have provided the best available evidence that eye cues make us feel watched. Across two experiments their results even suggest that those most sensitive to being watched are the most responsive to eye cues. In the first experiment, participants read a scenario in which they were asked to imagine wearing an embarrassing T-shirt while walking in a busy corridor with 30 people present or on a train with 50 people in the compartment. The participants were then asked to report how many people they thought would notice them. Results showed that those entering their estimation on a page with eye cues in the header reported significantly more people would notice them than those in a ‘no eyes’ control. Eye cues seemed to make people feel more watched.

In the second experiment the results suggested that participants’ sensitivity to being watched predicted their responsiveness to eye cues. Measuring participants’ sensitivity to being watched on the public self-awareness scale (Fenigstein, Scheier, and Buss, 1975; completed online via Amazon Mechanical Turk) their results indicated that those with very high or chronic self-awareness scores were more likely to donate some of their participant payment to a charity in the eye cues condition than in the control, whereas those with low public self-awareness showed no difference between conditions.

To summarize, then, Keller and Pfattheicher’s findings suggest that it is by making us feel watched that eye cues affect behaviour. First, because eye cues make people feel a higher state of public self-awareness, that is more aware of how they are presenting themselves to others. Second, because people’s differing trait levels of public self-awareness predict their responsiveness to eye cues, with those more sensitive to how they present themselves to others more strongly affected by eye cues than those who are less aware of the impression they create on others. If this is true, then we can expect eye cues to have similar effects on our behaviour to the presence of other people.

There is a large body of evidence showing that the ‘mere presence’ of others can affect behaviour (Markus, 1978; Schmitt et al., 1986), that gaining cooperative benefits depends often on sustaining a good reputation (Barclay, 2013; Izuma, 2012; Van Vugt, Roberts, & Hardy, 2007), and that reducing anonymity can make people more cooperative as it emphasizes reputational risk (Yoeli, Hoffman, Rand, & Nowak, 2013). Thus, if eye cues make people feel watched and thereby affect decisions and change behaviour, it is likely that they do so by prompting people to consider their reputation.

Consequently, we suggest that Northover et al. (2017a) may have found no effect of eye cues on generosity because whether one should give more or less when watched is not clear cut. Nettle et al. found in their (2013) dictator game that when eye cues were present participants were more likely to give something than nothing, but less likely to give a lot than those in the control condition. The eye cue effect seemed to be to reduce variation in donations, not to promote prosociality per se (Nettle, Harper, et al., 2013).

Similar effects have been found in a taking game (an inverse dictator game). Participants who took a lot in the control condition took less when eye cues were present, whereas participants who took very little in a control condition took more in the presence of eye cues (Chowdhury et al., 2014). Thus, although Northover et al. (2017a) showed there is no consistent shift in the proportion who gave when eye cues were present, it does seem that eye cues can both increase and reduce generosity dependent on the situation, suggesting that perhaps we should not expect consistent eye cue effects on generosity.

Such a deduction accords with the logic that being too conspicuously generous is not always reputation-enhancing (Klein & Epley, 2014; Klein et al., 2015) and, we contend, perhaps explains why the effects across the different experimental designs analysed in Northover et al. (2017a) reflect conflicting patterns and a neutral overall effect size.

Our meta-analysis suggests watching eyes produce a more consistent and robust effect on antisocial behaviour – the result, we suggest, of there being few situations, regardless of individual differences, in which being observed engaging in antisocial behaviour might be seen to be advantageous.

## 3.5. Limitations

Our meta-analytic model incorporates varied experimental designs which may weaken the overall findings. Stronger evidence for the eye cue effect on antisocial behaviour would have been forthcoming had there been more replications of successful experiments and less design diversity. Our meta-analysis also rests on narrow foundations and it would be premature to place too much weight on our findings without further research.

Additionally, we might be seeing a typical spate of strong effects in early research (de Winter & Happee, 2013; Ioannidis & Trikalinos, 2005; Monsarrat & Vergnes, 2018; Moonesinghe, Khoury, & Janssens, 2007), in this case in the eye cues and anti-social surveillance field, and thus an exaggerated effect size.

Finally, in looking only at the main effect of eye cues we have coded data as having shown no effect in Keep Britain Tidy's 2014 experiment when the authors did indeed report an effect in more nuanced analyses accounting for mediating and moderating variables. As more studies are published examining the watching eyes effect on antisocial behaviour, more nuanced meta-analysis will be possible providing greater insight than we are able to at present.

## 3.6. Further Research

Based on our findings, further laboratory and field studies designed to test the *watching eyes* effect on antisocial behaviour are clearly needed. Furthermore, there is a need to demonstrate that eye cues *do* make people feel more watched. Pfattheicher and Keller's (2015) finding that public self-awareness may moderate any effect of eye cues is a useful step in this direction. A replication of this finding would be a useful first step towards establishing how eye cues influence behaviour – if indeed they do.

## 3.7. Conclusions

Our meta-analysis of the eye cue effect on antisocial behaviour suggests eye cues may reduce the risk of antisocial behaviour by 35%. What matters here is not whether an effect size is large or small according to statistical guidelines but rather whether the effect size is meaningful (Aberson, 2011; Coe, 2002). In criminological analysis, for example, even smaller effect sizes may suggest significant reductions in crime (Weisburd, Telep, Hinkle, & Eck, 2010).

Our effect size, a 35% risk reduction for antisocial behaviour when eye cues are present, represents a meaningful effect in real-world terms. In its last *Costs of Crime* report the UK Home Office estimated that crime costs the UK economy between £35bn and £60bn a year (Brand & Price, 2000). Taking the lower end of the estimate, a 1% reduction in crime might be said to equate to a saving of £350,000,000. While we do not claim the watching eyes effect is anything like a panacea, the evidence so far suggests it could be a highly cost-effective criminal deterrent in some circumstances.

# Chapter 4

# 4

# Studies 2-4: Perceived Surveillance & Malicious Behaviour in the Joy of Destruction Game

> *"Under observation, we act less free, which means we effectively are less free."*
>
> Edward Snowden, Moscow, 2014

In Chapter 3, we found that the presence of watching eyes reduced anti-social behaviour by 35%. Here we tested this finding experimentally seeking to replicate and extend a previous study that had shown that the presence of eye cues reduced anti-social behaviour in the Joy of Destruction mini-game. We sought also to replicate findings showing that the presence of eye cues increased public self-consciousness scores. We extended the design to include a camera condition and tested for the possible moderating variables previously found to influence the direction and effectiveness of eye cues: religiosity, social value orientation and participants' independent vs interdependent self-construal. Three experiments were conducted: (1) a large-scale laboratory-based replication, finding no effect of eye cues or cameras on destruction rates or public self-consciousness; (2) an identical online replication with similar results; and, (3) an online study that attempted to reduce the psychological distance between participants by removing the role of chance in obscuring participants' decision to destroy – again finding no effect of either eye cues or the presence of a camera on destruction rates.

# 4. Studies 2-4: Perceived Surveillance & Malicious Behaviour in the Joy of Destruction Game

## 4.1. Introduction

The meta-analysis reported in Chapter 3 showed a 35% reduction in anti-social behaviour when eye cues were present in comparison with a ‘no eye cue’ control condition. The aim of the three studies reported in this chapter was to test these meta-analytic findings experimentally. We again added a camera condition to test whether the shape and design of the human eye was important for cueing any effect. To test the assumption that any effect of eye cues or cameras is due to them both promoting a sense of being watched we looked for differences in the extent to which participants’ public self-consciousness varied by condition. We then examined individual differences in self-construal, social value orientation and religiosity as possible moderators of any effect of watching eyes on anti-social behaviour.

Anti-social behaviour is defined broadly by the Oxford English dictionary as ‘*Contrary to the laws and customs of society, in a way that causes annoyance and disapproval in others*’. This definition leaves us with too subjective a construct to usefully test experimentally – laws and customs can and do change, annoyance and disapproval are subjective: what annoys *you* might not annoy me.

The fifth edition of the American Psychiatric Association’s Diagnostic and Statistical Manual of Mental Health Disorders (2013) offers a more helpful definition, explaining that anti-social personality disorder is ‘*characterised by a history of disregarding others and violating others rights…*’ with key elements including ‘*…deceit, manipulation of others, and failure to adhere to social norms.*’ But deceit and manipulation are not always anti-social. White lies can protect others from unhelpful or damaging truths that they have no ability to change. Similarly manipulating another can sometimes be in the manipulated individual’s self-interest and be socially approved. Social norms change over time and from group to group. We need something more precise if we are to examine it effectively.

We suggest that anti-social behaviour defies precise definition as it is a relational concept, dependent on cultural, social and temporal context. For example, stealing might be anti-social, unless the person stealing was starving, impoverished and or/stole in order to feed a starving child, in which case we might even regard the theft as a prosocial act. Terrorists and their supporters often regard acts such as the killing of large numbers of civilians as profoundly pro-social, even altruistic (O'Gorman & Silke, 2015), while for others such acts are the apogee of anti-sociality.

Although the terms prosocial and antisocial are used widely in the psychological literature (e.g. Bateson, Callow, Holmes, Roche, & Nettle, 2013; Oda, Kato, & Hiraishi, 2015) we believe that these generic terms are unhelpful. It would be desirable to have a more precise concept to examine. Evolutionary biologists use the term 'spiteful' describing spiteful behaviours as those where an organism 'harms itself in order to harm another individual more' but finds such behaviours to be exceedingly rare and principally found in species in decline and thus, over time, eliminated by natural selection (Hamilton, 1970), although low-levels of spite might be an evolutionarily stable strategy in small populations (Knowlton & Parker, 1979) or perhaps in group selection terms (Wilson, 1975), as explicable as advantageous to the species or population at the expense of the individual organism . In human behaviour harming oneself to benefit another individual or group might be considered not spiteful but altruistic, again as in terrorist attacks (O'Gorman & Silke, 2015). In their widely cited paper on spiteful behaviour economists Pillutla and Murnighan (1996) use the word spite as '*the behavioural reaction that accompanies anger and which is designed to hurt the offending other*'. In this conception spiteful relates to inequality aversion, or justice, and thus can be seen as a prosocial corrective in the interests of the group. More recent work by economists Kimbrough and Reiss construes a spiteful act to as one that ‘directly imposes harm on another and provides no immediate benefit to the spiteful actor’ (2012) which better describes the behaviours we are interested in. However, the wider definitional confusion surrounding the word ‘spite’ is best avoided.

The Oxford English definition of spiteful “showing or caused by malice” suggests an alternative in the term ‘malicious’, defined by the OED as *the desire to harm someone; ill-will*. Indeed, the term is sufficiently precise to carry a legal and ethical dimension where it is defined as *Wrongful intention, especially increasing the guilt of some offences*. Therefore, we seek to frame anti-social in this context as synonymous with malicious and its definition as *the desire to harm someone*. By choosing this restrictive and more precise term we are able to ensure our experimental design is testing for precisely the effect we are interested in, leaving less room for ambiguity and challenge.

### 4.1.1. Joy of Destruction Task

We selected the Joy of Destruction (JoD) game to examine the effect of eye cues and cameras on malicious behaviour. In the JoD game two participants are allocated X amount of money. Both participants are anonymous and unaware of each other's identity. Both are asked independently if they would be willing to pay X amount to destroy X amount of the other subject's endowment. Neither know the other participant’s decisions when making their own. There is a cost to choosing to destroy, but no possible benefit – in all scenarios the participant is worse off for having chosen to pay to reduce the other’s allocation, regardless of what the other participant may choose. We judge that this behaviour meets our definition of malicious behaviour – the intention must be to hurt the other and is sufficiently strong that a participant is willing to pay to inflict this pain. The task has been run with two conditions (Abbink & Herrmann, 2011; Abbink & Sadrieh, 2009). A ‘hidden’ condition in which there is a one-in-three chance a participant’s allocation will be reduced regardless of what their ‘partner’ decides, and an ‘open’ condition in which there is no role of chance in outcomes and so participants know that any destruction inflicted on them is the result of a wilful act from their anonymous partner. Baillon, Selim, and van Dolder (2013) used just the ‘hidden’ version of the task in their test of the effect of eye cues on destruction decisions in the JoD, while Ortmann (2013) used just the ‘open’ condition in their version designed to test the role of Machiavellian personality traits and the desire for influence in decision-making in the JoD.

Abbink and Sadrieh (2009), who first introduced the JoD task, used the open and hidden conditions to show that the degree of psychological distance between ‘partners’ affected the frequency of decisions to destroy. Participants were mutually anonymous, but the psychological distance between them was reduced in the open condition – since each would know that any reduction in their income had been inflicted on them by the other and not by chance. When this was the case, the overall frequencies of destruction were low. Participants destroyed some of the other participant’s allocation in just 8.5% of cases (n = 20) across 10 rounds of the task. However, when the participants were told that there was a random chance that both their and the other participant’s earnings might be destroyed by chance (‘hidden’ condition – because a participant could not know why their allocation had been reduced), the wilful destruction rate was 39.4% (n =20). The difference between the open and hidden conditions was significant for all rounds.

In a replication of this experiment, Abbink & Hermann (2009) required participants to make decisions as to whether or not to destroy or not in a one-shot game, again comparing open and hidden treatments. In their replication, destruction rates were again significantly different: low in the open condition (n=69) where 10.1% of participants elected to destroy the other’s income, but high in the hidden condition where 25.8% of participants chose to destroy (n=62). All those who have employed the Joy of Destruction task agree that the increased destruction in the hidden condition is likely due to the increased psychological distance between destroyer and ‘victim’ (Abbink & Sadrieh, 2009; Abbink & Hermann, 2009). In the open condition the ‘victim’ will know that another person was responsible for their reduced income, and may ruminate on this. The destroyer may therefore be unwilling to choose to make the victim suffer. Whereas in the hidden condition the victim could just consider themselves unlucky. The destroyer may then ‘take pleasure in being nasty’ while knowing that their decisions to destroy will inflict little mental suffering.

There is support for this idea from other studies. Aimone & Houser (2012) describe and demonstrate the powerful effects of ‘betrayal aversion’ in an economic trust game. There were two versions of the task in their study that are relevent to our research. In the first version of the task, trusting the other participant could earn you more but risked them betraying you to maximise their income, leaving you earning much less. As an alternative strategy, you could not trust them and both earn a small income – but less than if you had both co-operated. In this first version of the task, you knew if you were betrayed by your anonymous partner – in effect an ‘open condition’ similar to that described in the JoD. In the second version, you decided whether to trust or not, but all betrayals were determined by the computer – all decisions were ‘hidden’. Aimone and Houser showed that participants trusted most when they would never know if they had been betrayed or not – trusting 92% of the time in the hidden condition compared to 65% in the open. Other studies have found similar evidence for such ‘betrayal aversion’ in a similar task, globally across 8 different societies (Bohnet, Greig, Herrmann, & Zeckhauser, 2008), and in an ultimatum game in the US where participants could ‘hide’ their decision behind a coin toss (Cox & Li, 2012).

Baillon, Selim and Van Dolder (2013) also employed the JoD game, again as a one-shot task, to test the effect of watching eyes on anti-social behaviour. All participants and decisions were anonymous and there was a 1/3 chance that participants’ allocations would be destroyed irrespective of their partner’s decision to destroy or not. Participants were randomly allocated to ‘eyes’, ‘no eyes’ and a ‘peers’ condition with decisions ‘hidden’ by the 1/3 random chance of destruction in all three conditions. Where previous studies (Abbink & Sadrieh, 2009; Abbink & Hermann, 2009) ran experiments as pencil and paper exercises, Baillon et al (2013) employed the JoD as an online computer-based task. Their design also differed from that used in most eye cue experiments in that the statue of Erasmus was used rather than the more frequently used stylised or human eye images (See Figure 4.1.). The peers condition used photographs of groups of students in university scenes.

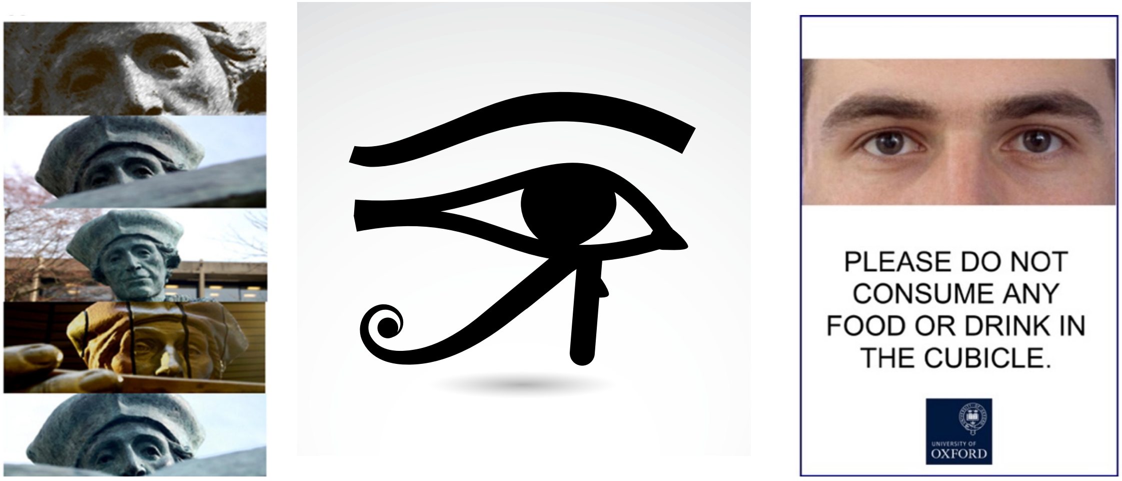

PLEASE DO NOT CONSUME ANY FOOD OR DRINK IN THE CUBICLE.
UNIVERSITY OF OXFORD

*Figure 4.1. Left to Right: Erasmus statue as used by Baillon et al (2013), stylised eye of Horus design used by Haley and Fessler (2005), eye posters adapted from that used in Fathi, Bateson, and Nettle (2014).*

Despite these differences in design, Baillon et al (2013) found similar results to Abbink & Hermann (2009) with 24.84% of the participants deciding to destroy ($N$=153), compared with 25.8% in Abbink and Herrmann's (2009) experiment. However, Baillon et al found that destruction rates were much reduced in the eyes (N=49) and peers (N=53) conditions when compared with the control ($N$=51). The differences were significant (control 25.8%; eyes 17.65%, $\chi^2(1)= 5.534$, $p = .02$ peers: 18.87%, $\chi^2(1) = 4.96$, $p = .03$) with no difference between the eyes and peers conditions ($\chi^2(1)=0.03$, $p=0.87$).

We chose the JoD as our task for two reasons. Firstly, as a pure measure of anti-social – malicious - behaviour because the intention behind destroying can only be to hurt the anonymous partner. Secondly, because it enabled us to seek to replicate the findings of Baillon et al.'s (2013) watching eyes study. This attempt was useful and justified as the results of Baillon et al. had contributed to our meta-analytic findings and because there exists a growing scepticism concerning the replicability of many eye cue experiments (Northover et al., 2017a, 2017b). We planned from the outset to analyse not just differences in mean destruction by condition, but also whether there was any difference in the proportion of people who never chose to destroy and those who did so at least once. Nettle, Harper, et al. (2013) found that while eye cues didn't always alter mean donations in dictator games, they did more reliably alter the proportion of participants choosing to give something rather than nothing, while reducing the number giving everything away. Nettle et al. suggested this might be due to people avoiding a reputation either for being ungenerous or for being foolishly over-generous. We hypothesised that the proportion choosing never to destroy ought to decrease in the presence of watching eyes in order to avoid appearing naïve. Similarly, the proportion choosing to destroy all the time ought to decrease as people are cued to avoid a reputation for wilful malicious behaviour.

### 4.1.2. Public Self-Consciousness

In our experiment on surveillance and prosociality (Chapter 2) we measured the extent to which people felt watched indirectly by employing the Snyder Self-Monitoring Scale to examine whether people self-monitored more in the presence of eye cues and whether high-self monitors were more sensitive to eye cues. There was no evidence that eye cues or the presence of a camera were enough to make people feel watched in that study. In studies 2, 3 & 4 reported in this chapter we elected to use the same public self-consciousness scale as Pfattheicher and Keller (2015) in an attempt, firstly, to more directly measure the feeling of being watched than in our first experiment, secondly, to replicate Pfattheicher and Keller's findings and thirdly to provide empirical justification for the explanation that watching eyes reduce anti-social behaviour by making people feel watched, thereby heightening their reputational concerns.

We also examined responsiveness to surveillance cues across three individual difference measures. (1) Independence-Interdependence scores, (2) Social Value Orientation (pro-self vs pro-social), and (3) religiosity. We next explain the selection of these choices in turn.

### 4.1.3. Independence-interdependence Scores

Kitayama et al. (2004) demonstrated that there are cross-cultural differences in responsiveness to eye cues – in their case via stylised cartoon faces – and showed that the differences are best explained with reference to differences in identity formation between cultures: some have strong collective identities, their identity is *interdependent*, others are highly individualised, and identity is said to be *independent*. In Kitayama et al.'s study Japanese, and Asian Americans were asked to rate the desirability of CDs, ranking their top 10 from an extended list. Later they were asked to select one CD to keep, choosing between their 5th and 6th favourites, and then re-rate the desirability of the CDs. Japanese and Asian Americans kept their ratings stable when there was no social cue present. When the social cue was present they give a higher rating for the CD they chose and a lower rating for that they rejected. This greater choice-justification or spread of alternatives in the presence of eye cues was not seen in Caucasian and Black Americans. Kityama et al. suggest this is because Japanese and Asians feel a greater identity threat when they feel others are evaluating their decisions. As a result, they feel the need to adjust their ratings to justify their decision to themselves, feel more comfortable with the decision and thus less threatened. In contrast, black and white Americans are not affected by eye cues. They don't fear the threat of social sanction. When asked to re-evaluate their decisions they felt their competence and efficacy threatened by self-judgement if the decision seemed arbitrary and thus they consistently rated the selected CD higher and the rejected one lower in both social cue and control conditions than Japanese and Asians, with no difference in the presence of eye cues. Kitayama et al. conclude that Japanese and Asians feel more cognitive dissonance, and more identity threat, from the perceived gaze of others, and are thus more influenced by eye cues. Whereas, they suggest, black and white Americans are interestedly primarily only in their own judgement of themselves and are thus not influenced by the presence of eye cues.

The difference that Kityama et al. found between the Japanese and Asian Americans in comparison with Caucasian and Black Americans was explained largely by the more interdependent culture of the Japanese and of Asian Americans. This greater interdependence is well attested to in the literature (for example: (Geertz, 1975; Hamedani, Markus, & Fu, 2013; Heine, 2001; Kitayama et al., 2014a; Kitayama, Park, Sevincer, Karasawa, & Uskul, 2009; H. R. Markus & Kitayama, 1998; Triandis, 1989) (though see Matsumoto (1999) for a dissenting view). Asians' greater responsiveness to social cues - in this case a stylized cartoon face - may reflect the overlap between individual and group identity. When your identity is more dependent both on the collective group reputation, and the group's judgement of your actions, being cued to feel watched cues an acute identity threat. In contrast, more independent individuals worry about their self-esteem as judged internally and so are less concerned by being watched (Imada & Kitayama, 2010; Kitayama et al., 2004). Drawing on this understanding, Kitayama and colleagues have shown that individual differences in interdependent and independent self-construal influence the direction of responses to social cues within cultures too (Imada & Kitayama, 2010; Kitayama et al., 2004; Na & Kitayama, 2012; Park & Kitayama, 2012). Consequently, we measured interdependence and independence scores in order to understand whether participants' responsiveness to eye cues or the presence of a camera was moderated in degree or direction by the extent to which their identity was dependent on the judgement of others. We predicted that more interdependent participants would be acutely sensitive to reputational cues, eager to protect their reputation ‘in the eyes of group’ and so might be less likely to destroy in the presence of eye cues or a camera than in a control condition, whereas low interdependent participants’ destructions rates would not vary by condition. Beyond the theoretic importance of this measure, differences in response to surveillance across cultural would be of profound consequence and interest in policy and business circles. For example, they might affect the formulation of international legal norms, the suitability of employing surveillance in peacekeeping or in business and financial regulatory practice, or in designing information architectures in smart devices and web-based systems.

### 4.1.4. Social Value Orientation (SVO)

For our second individual difference measure we selected Social Value Orientation (SVO) as another variable that has been found to predict greater sensitivity to eye cues in previous research. SVO divides people into largely self-explanatory pro-self and pro-social categories (Van Lange, 1999; Vanlange, Otten, Debruin, & Joireman, 1997). Luo, Zhang, Tao, and Geng (2016) conducted two experiments employing anonymous Prisoners’ Dilemma games to test responses to eye cues and examine any interaction effects with individuals’ SVO and found that participants were more likely to cooperate in the presence of both subliminal and supraliminal gaze cues (faces) if they were prosocial/high SVO. Pro-selfs were unaffected by the presence of eyes in the experiment testing supraliminal cues, and less cooperative after being exposed to subliminal cues.

Consequently, we included an SVO questionnaire to understand our participants’ SVO ratings and any interaction effect with frequency of destruction. We predicted that (a) high SVOs, pro-socials, would destroy rarely in all three conditions (control, eyes and camera) since their motivation is to do well by others. (b) Low SVOs, pro-selfs, should be more sensitive to cues of social surveillance destroying less often in the presence of eye cues or a camera in order to protect their reputation and maximise future pay-offs, whereas (c) in the absence of surveillance cues pro-selfs may destroy more to hurt others and to prevent others hurting them unanswered.

Our third SVO prediction, that pro-selfs might be guided to expect the worst of others and so act to ‘even the score’, is supported by Irlenbusch and Ter Meer (2013) who showed that in anonymous public goods games people judge what their partners will do by what they themselves would do in the same situation. Low SVOs have low expectations of themselves and therefore of others, assume that they can’t be trusted and so cooperate less. Again, beyond the theoretic value of SVO, any moderating affect would be important in public policy helping plan for the fielding of surveillance measures only where the population targeted is comprised of individuals likely to be affected by the intervention. Thus, if terrorists are highly pro-social, but fraudsters highly pro-self, as seems possible, the evidence from previous research (Luo et al., 2016) predicts that surveillance measures might be used again the fraudsters effectively, but perhaps shouldn’t be expected to have much effect on terrorists.

### 4.1.5. Religiosity

Religiosity may also be related to individuals’ responsiveness to feeling watched. Punishment in social groups is necessary for group success, preventing free-riding and social infractions, and enforcing hierarchy to enable group functioning and ensure group cohesion (Fehr & Gachter, 2002; Gürerk, Irlenbusch, & Rockenbach, 2006). But as group size increases, it becomes harder to track individual reputations, providing greater opportunity for free-riding and selfish behaviours to go unpunished. The supernatural surveillance hypothesis argues that humans may have evolved religion as a novel response to this challenge, developing omnipresent moralizing spirits and gods in order to ensure compliance to social norms as human group size increased (See Chapter 1). We predict therefore that highly religious individuals would be less affected by the presence of eye cues or the presence of a camera than those without religious belief since believers would already feel watched by an omnipresent moralizing god even in the control condition. To examine whether this is the case, participants were asked to complete the ten-item variant of the Centrality of Religion Survey (Huber & Huber, 2012) to assess how great a role religion played in their lives, as a proxy for how much they felt watched by God.

### 4.1.6. Format & Overview

Three studies were conducted. Study 2 was a large-scale laboratory study which sought to replicate and extend Baillon et al. (2013); Study 3 aimed to replicate Study 2 online. Study 4 was a repeat of Study 3 online but the experimental design was modified to remove the role of chance in 'hiding' participant decisions in order to reduce the psychological distance between decision makers. Finding no major differences between destruction rates across the 3 studies justified analysing the data in the aggregate and this analysis is reported after the 3 experiments. I felt it necessary to comment on the significant interactions at the end of each study for the sake of thoroughness and to show how the studies linked together (or not). Some readers may prefer to skip the individual discussion sections to focus on the main conclusions drawn in the final discussion at paragraph 4.6. on p. 232. All graphs and regression tables can be found side-by-side for each variable at Appendix B.

## 4.2. Study 2

Study 2 was a laboratory-based study which examined the effect of watching eyes or a camera on participants malicious behaviour in the Joy of Destruction task.

### 4.2.1. Method

#### 4.2.1.1. Participants

A priori power calculations in *R* (R Development Core Team, 2016) and G*Power (Faul et al., 2007) (see additional materials) suggested 199 participants were needed to achieve sufficient statistical power.

Recruitment was undertaken in advance of the study though additional participants did sign-up at most locations on our arrival. Participants were anonymous at all times. All participants were paid a £4.00 show-up fee.

Participants in Study 2 were a mixture of civilian staff and military personnel from the Royal Navy, Army and Royal Air Force, and predominantly male. Given that previous research has found little evidence for any influence of gender on the Watching Eyes Effect, no attempt was made to balance groups by gender.

Our sample in Study 1 was 67% male (*N=137)*, 31% female (*N* = 64), 2% (*N*=4) did not report their gender. Ethnically, our sample was 95% white or white British. We did not collect age data in Study 1, an oversight corrected in Studies 3 & 4.

#### 4.2.1.2. Design

Participants first completed the Joy of Destruction task which measured how often participants paid to reduce another participant's income. Second, they completed the Social Value Orientation 'slider' deciding how much of a hypothetical amount of money they would allocate to themselves or others. Third they completed the questionnaire measures on Public Self-Consciousness, Self-Construal and Centrality of Religion as detailed below.

The independent variable was surveillance condition, within which there were 3 groups: a no-surveillance control, a watching eyes condition, and a camera condition. The primary outcome variable was the frequency of destruction. The second outcome variable was public self-awareness. The studies also explored public self-consciousness as a potential co-variate and three additional potential co-variates: Social Value Orientation, Self-Construal and Centrality of Religion.

#### 4.2.1.3. Materials

All questionnaires were presented using Limesurvey (Limesurvey GmbH., 2003).

##### 4.2.1.3.1. Public Self-Consciousness Scale

Participants completed the public self-consciousness scale (Fenigstein et al., 1975), as revised by (Scheier & Carver, 1985). This 22-item questionnaire assesses private and public self-consciousness, as well as social anxiety. Participants respond to statements such as ‘I’m self-conscious about the way I look’, ‘Large groups make me nervous’, or ‘I usually worry about making a good impression’ on a 4-point scale from 0 = not like me at all, to 3 = a lot like me. Participant’s scores were the sum of the 22 responses, with two items reverse coded.

It is only public self-consciousness that we are interested in and report in this study. Based on the findings of Pfattheicher and Keller (2015) we employ it to test for whether watching eyes made participants feel more sensitive to how they were presenting themselves to an audience, a proxy for directly testing whether eye cues make people feel more watched.

##### 4.2.1.3.2. Singelis Self-Construal Scale

The Singelis Self-Construal Scale was used to gather interdependence scores based on a 30-item questionnaire, 15 relating to how independent a participant’s self-construal was, and 15 relating to how interdependent they saw themselves to be. Participants were asked to read statements such as ‘I enjoy being unique and different from others in many respects’, ‘I feel good when I cooperate with others’, ‘I will stay in a group if they need me, even when I am not happy with the group' and respond on a 7-point Lickert-type scale from 1 = Strongly Disagree to 7 = Strongly Agree. Participants' independence scores were calculated by summing participant responses (1 to 7) for the 15 items related to independence and dividing by 15 to give the mean response score. Interdependence scores were calculated in the same way. We were interested in the continuous independence-to-interdependence scale, found by subtracting participants' independence score from their interdependence score. (Kitayama et al., 2014b; Singelis, 1994).

#### 4.2.1.3.3. Social Value Orientation

Social Value Orientation was assessed using the 6-item SVO slider measure which scores individuals on a scale from prosocial to pro-self (with extremes of 'altruistic' and 'competitive' at either end of the scale). The scale is based on their choices of how participants would hypothetically divide set-amounts of cash between themselves and others, as devised by Murphy, Ackermann, and Handgraaf (2011). Scores were calculated by (a) deriving the mean of the amount allocated to self ($M_s$) and the mean allocated to others ($M_o$) in a subject's responses (b) subtracting 50 from both means (c) calculating the inverse tangent of the mean of the payoffs allocated to the other minus 50 and the mean of the payoffs to the self minus 50:

$$SVO^{\circ} = arctan\left(\frac{M_0 - 50}{M_s - 50}\right)$$

Higher scores indicated higher prosociality, and we employed this measure as a continuous variable. Thresholds for calculating the categories were employed only to better understand our sample: Competitive = SVO° < –12.04°, Individualist = SVO >-12.04°, <22.45°, Prosocial = SVO >22.45°, <57.15°, Altruist = SVO° > 57.15°.

### 4.2.1.3.4. Centrality of Religion Scale

Participants' religiosity was assessed using the ten-item variant of the Centrality of Religion Scale (CRS), which scores participants according to the degree to which religion features in their lives (Huber & Huber, 2012). Participants were asked questions such as 'How often do you pray?' and 'To what extent do you believe in an afterlife? e.g. immortality of the soul, resurrection of the dead or reincarnation?'. They were asked to respond on one of three multiple-choice scales as described in Table 1 & 2 below. Centrality of religion scores were calculated by summing participant responses and dividing by 10 to find the mean ranging between 1.0 (not religious) and 5.0 (highly religious). Tables 1 and 2 display the scoring system for the CRS questionnaire.

Table 4.1. Scoring for the Centrality of Religion Score, 1 of 2.

| Objective frequencies of prayer (personal and obligatory) and meditation | Recoding into five levels | Objective frequencies of participation in religious services | Recoding into five levels |
|---|---|---|---|
| A) Several times a day | 5 | A) More than once a week | 5 |
| B) Once a day | | B) Once a week | |
| C) More than once a week | 4 | C) One or three times a month | 4 |
| D) Once a week | 3 | D) A few times a year | 3 |
| E) One or three times a month | | E) Less often | 2 |
| F) A few times a year | 2 | F) Never | 1 |
| G) Less often | | | |
| H) Never | 1 | | |

Table 4.2. Scoring for the Centrality of Religion Score, 2 of 2

| | Score | Frequency | Importance |
|---|---|---|---|
| Categories of a five-level answer-scale | 5 | very often | very much so |
| | 4 | often | quite a bit |
| | 3 | occasionally | moderately |
| | 2 | rarely | not very much |
| | 1 | never | not at all |

### 4.2.1.4. Stimuli

Participants were exposed to one of three conditions in the experiment: Control, Eye Cues, and Camera (see Figure 4.2). As with our previous experiment (Chapter 2) examining the effect of watching eyes on generosity, we selected the ‘watching eyes’ and control posters used by Nettle, Harper, Kidson, Stone, Penton-Voak and Bateson (2013) as the eye cue in our experiment since they had found that the presence of these cues led to an increased probability of donating something rather than nothing (80% donated in the ‘eyes’ condition compared to 63% in the control). We modified these by replacing the Newcastle University logos with that of Oxford and, for the control posters, The Institute of Neuroscience Logo with the crest of Magdalen College, Oxford. Four posters were placed in the booth (see *4.2.1.5 Procedure* & Figure 4.2 below). In the control and camera condition the posters with the Magdalen crest was displayed. These were replaced with the eyes posters for the eyes condition. In addition, the poster design and text appeared on-screen below the task – control image in the control and camera condition, eyes image in the eyes condition.

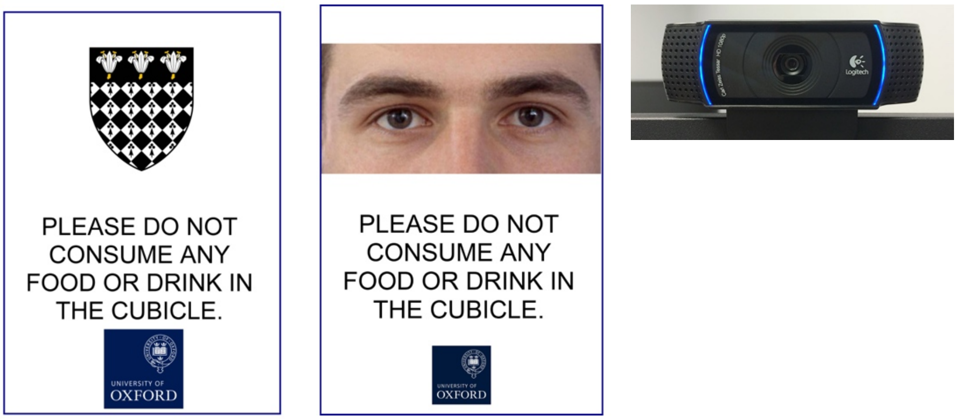

PLEASE DO NOT CONSUME ANY FOOD OR DRINK IN THE CUBICLE.
UNIVERSITY OF OXFORD
PLEASE DO NOT CONSUME ANY FOOD OR DRINK IN THE CUBICLE.
UNIVERSITY OF OXFORD
Logitech

*Figure 4.2. (From left to right) Condition 1: Control; 2: Eye Cues; 3: Camera.*

For the camera conditions we selected the Logitech C920 HD Pro USB 1080P Webcam. The camera was switched on, which meant that a blue light illuminated on the camera, drawing attention to its presence. Participants in the camera condition were told, via the on-screen consent form, that a camera would be present and had to grant permission to be filmed in each round to ensure they were always aware that the camera was present. Any participants that were assigned to the camera condition but did not approve the use of the camera were switched to the control condition since in this scenario they were functionally the same.

#### 4.2.1.5. Procedure

Participants were tested under laboratory conditions in separate rooms at six locations: Royal Air Force bases High Wycombe, Digby and Odiham, the Defence Academy (UK), MoD Main Building (Whitehall), and Denison Barracks. Testing rooms varied in size but were uniformly plain-walled, white-washed with strip-lighting. Any windows were temporarily papered over. Participants were anonymous at all times once in the testing room.

All participants completed the study via a desktop computer in an Allermuir Haven pod (see Figure 3). Participants were briefed, gave their consent, completed the study and were debriefed at the desk by written on-screen instructions and forms via our online platform code, which is available from the Open Science Framework for examination and for use by future researchers (Dear & Ellen, 2018).

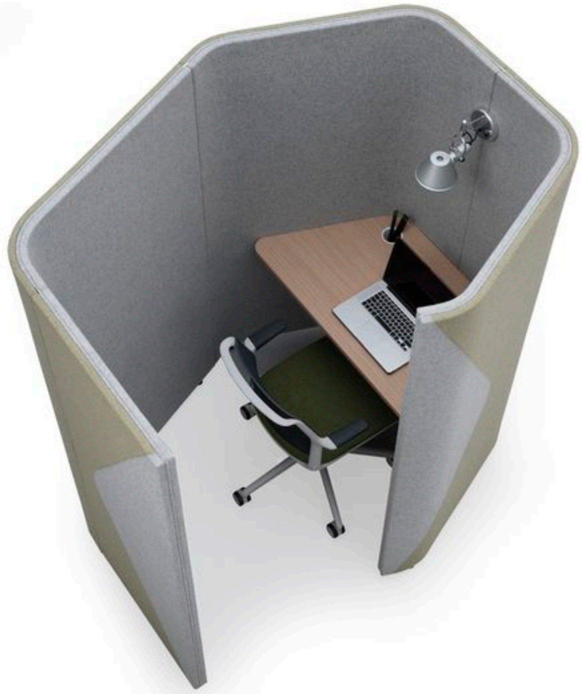


*Figure 4.3. Schematic of Testing Pod Used in the Study*

The lead experimenter was always the same individual (KD), although research assistants varied at each location. Conditions were arranged and re-arranged by the research assistant who was unaware of the purpose of the study and had no access to the data collected. The role of the assistant was to escort participants to and from the booth, and to ensure participants in different booths and rooms did not meet. All data was collected via the web-based Joy of Destruction task.

In our version of the Joy of Destruction task participants completed ten rounds of a task, believing they were paired with a different anonymous partner in each round. Participants were allocated £1.00 per round and could pay 10p (£0.10) to reduce their partner's allocation by 50p (£0.50), while their partner could concurrently do the same. Participants knew that there was a random 1/3 chance that their own income or that of the other participants they were paired with in each round would be reduced regardless of their or the other participant's decision. They also knew that there was no way of either party knowing if any reduction was a result of chance or the other's decision. Participants knew nothing of the other participants decisions until after all ten rounds of the task were complete.

As mentioned, participants believed they were paired anonymously, over the internet, with a different anonymous partner in each of ten rounds. Thus, all interactions were one-shot and independent. In reality, because of the difficulty of keeping sufficient numbers of partners logged in to the experiment at the same time, the first participants were paired with ten randomly generated decisions and later participants' ten decisions were matched with ten random decisions of other participants from previous iterations of the task. Participants were informed of this deception in a post-experiment debrief form and offered the chance to withdraw from the experiment should they wish to. No participants chose to do so.

#### 4.2.1.6. Data Plan

Preliminary analysis – calculation of individual difference scores, mean destruction, and mean expected destruction and removal of incomplete or corrupted responses was conducted in Excel. All further data exploration, transformation and analysis was undertaken in *R* (R Development Core Team, 2016).

We originally planned to conduct the parametric tests described in this section. However, since our data failed the assumptions check (see section 4.2.2.) we had to adapt our analysis plan to incorporate the non-parametric distribution of the results obtained.

To test for the effect of our independent variable condition (Control, Eyes and Camera) on our dependent variable mean destruction in the JoD we planned to conducted a one-way ANOVA. Additionally, since we expected that the majority of participants would choose not to destroy at all (zero of 10), we planned to transform our destruction data to the binomial 'destroyed or not' and test whether condition had any effect in changing the proportion who never destroyed in a 3 x 1 logistic regression, 'destroyed or not' by condition. We also planned an exploratory logistic regression analysis to examine whether surveillance cues might change the proportion of participants destroying every time (10 of 10) compared to those who destroyed less frequently.

To see if surveillance cues increased participants' self-awareness we planned to test the effect of condition (Control, Eyes and Camera) on Public Self-Consciousness scores in a one-way ANOVA.

To explore whether Public Self-Consciousness, Social Value Orientation, self-construal and religiosity modify any effect of condition on mean destruction we planned to add these as continuous variables (i.e. as covariates) to an ANOVA with condition as a factor and mean destruction as the dependent variable.

Data & *R* Code for Studies 2-4 is available at Open Science – see note at Appendix D.

### 4.2.2. Study 2 Results

The distribution of destruction rates in Study 1 violated the normality assumption, both in the aggregate, $W = 0.82$, $p < .001$ and by condition, Control $W = 0.81$, $p < .001$, Eyes $W = 0.85$, $p < .001$, Camera, $W = 0.79$, $p < .001$. Levene's test indicated unequal variances across condition, ($F$ (2, 207) = 3.27, $p = .04$). Histograms indicated the data was heavily right-skewed. Neither square root, cube root nor log or Tukey transformations were sufficient to enable parametric testing. Consequently, all analyses that include destruction rates were conducted using the appropriate non-parametric test. Public self-consciousness scores, independence-interdependence scores and religiosity scores were also not normally distributed.

Since our data for all variables except self-construal scores failed the assumptions checks we could not conduct our ANOVAs as planned. Consequently, we employed the Kruskal-Wallace test (Kruskal, 1952) in *R* using the method described by Field, Field, and Miles (2012, pp. 674-686) to test for any main effect of condition on all our non-parametrically distributed variables to provide the summary statistics at Table 4.3. The Kruskal-Wallace test is functionally equivalent to a one-way ANOVA, but tests differences in the distribution of ranked-summed data between groups. Consequently, we follow the convention of reporting the median and inter-quartile range (Field et al. 2012; p. 686). A one-way ANOVA *was* used to test for any differences in self-construal scores by condition since self-construal scores passed the assumptions checks.

Table 4.3 shows the results of a between-group analysis with watching eye condition (control, eye cue, camera) as between-subjects variable and destruction rates as the dependent variable. Kruskal-Wallace tests showed no significant differences between groups: median destruction rates did not differ significantly between the control and eye cue and camera conditions $H(2) = 0.96, p = .62$ Additionally, there was no significant difference in self-awareness scores by surveillance condition.

Again, adapting our data plan to accommodate our non-parametrically distributed results we employed a Poisson generalised linear regression model (GLM) to test for the influence our covariates (Crawley, 2012, pp. 381-385). Poisson GLMs are suitable for non-parametric data and particularly for analysis of count data (Field et al., 2012, p. 329) such as found in the primary dependent variable ‘destruction’ which was a count of the number of times a participant destroyed and could vary from 0-10.

A robust regression analysis (Field & Wilcox, 2017; Wilcox, 2017) was also undertaken as a confirmatory check on our Poisson GLM analysis but results did not differ and we do not report this for the sake of brevity.

## 4. Studies 2-4: Perceived Surveillance & Malicious Behaviour in the Joy of Destruction Game

Our Poisson regression analyses found no significant differences between our exploratory individual difference measures, the potential moderating variables Independence-Interdependence- scores, SVO scores, or religiosity scores across the surveillance conditions.

Table 4.3. Joy of Destruction Study 2 Results Stratified by Condition

| | | Control | | Eyes | | Camera | | Test Statistic | *p* | Test |
|---|---|---|---|---|---|---|---|---|---|---|
| *N* | | 71 | | 69 | | 66 | | | | |
| Male/Female/Unknown | | 49/20/2 | | 42/25/2 | | 46/19/1 | | | | |
| Destruction (0-1) | median [*IQR*] | 0.10 | [0.10, 0.50] | 0.10 | [0.00, 0.50] | 0.15 | [0.00, 0.47] | *H = 0.96* | .62 | K-W |
| Expected Destruction | median [*IQR*] | 0.40 | [0.00, 0.50] | 0.40 | [0.10, 0.50] | 0.30 | [0.00, 0.50] | *H = 0.21* | .90 | K-W |
| Public Self-Consciousness | median [*IQR*] | 10.00 | [8.00, 13.50] | 11.00 | [8.00, 15.00] | 11.00 | [7.00, 14.00] | *H = 0.82* | .66 | K-W |
| Independence-Interdependence | mean (*sd*) | 0.11 | (0.97) | 0.13 | (0.83) | 0.06 | (0.85) | *F = 0.11* | .90 | ANOVA |
| SVO Score | median [*IQR*] | 44.01 | [40.87, 45.00] | 44.00 | [43.31, 45.00] | 44.00 | [40.93, 45.00] | *H = 1.92* | .39 | K-W |
| SVO Cat | Prosocial (%) | 68 | (95.8) | 68 | (98.6) | 60 | (90.2) | | | |
| CRS Score | median [*IQR*] | 1.80 | [1.50, 2.55] | 1.90 | [1.40, 2.30] | 1.90 | [1.60, 2.40] | *H = 1.43* | .49 | K-W |

Values presented refer to the median, inter-quartile range, p-value and significance test associated with each variable. *H* = Kruskal-Wallis Chi Squared.

Over 50% of our participants chose not to destroy at all over the full ten rounds of the task. We transformed our data in to the binary variable 'destroyed or not' to differentiate between those never choosing to destroy and those choosing to destroy at least once. We wanted to see if a greater proportion chose never to destroy in the presence of eye cues or a camera than in the control condition. Logistic regression showed that our model was not significantly better than chance at predicting whether participants destroyed or not, $\chi^2(2)$ 1.61, $p < .44$. As we would expect, given the non-significance of the model there was no systematic variation in the probability of destruction by condition (Eyes $z = 1.22$, $p = 0.22$; Camera $z = 0.29$, $p = .78$). Table 4.4. displays the results of the logistic regression.

In accordance with the recommendations of Field et al. (2012, p. 341) we calculate and report three variations of $R^2$ throughout this chapter to check the goodness of fit. Hosmer-Lemeshow's $R^2$ was calculated by dividing the model chi-square by the deviance of the model before any predictors were entered. Hosmer-Lemeshow's $R^2$ can vary between 0 (indicating that the predictors do not predict the outcome variable) and 1 (indicating that the model predicts the outcome variable perfectly). In Study 1, Hosmer-Lemeshow's was very small $R^2 = .01$. An alternate measure Cox-Snell's $R^2$, which is also based on the deviance of the model and has a similar distribution but cannot reach its theoretical maximum of 1, was calculated as a confidence check and, after rounding, was the same, $R^2_{CS} = .01$. Nagelkerke's $R^2$, another alternative to Hosmer-Lemeshow which *can* reach its theoretical maximum and, after rounding, was the same, $R^2_N = .01$. For brevity, they are henceforth reported only in the table. The results of the logistic regression are displayed in Table 4.

We had intended to examine any differences between the proportion of those choosing to destroy 100% of the time by condition. However, too few participants chose to destroy every time to enable meaningful analysis. In fact, only 1 of 206 participants chose to destroy every time. We considered lowering the threshold but zero participants chose to destroy nine times of ten. Any further lowering of the threshold would have been arbitrary, post hoc and not sufficiently aligned with our initial hypothesis. Consequently, analysis of extreme destruction was not possible and was not carried out.

Table 4.4. Study 2: Logistic Regression on the binary variable Destroyed or Not Comparing Surveillance Cue Conditions Eyes and Camera with a no surveillance cue control.

| | | 95% *CI* for odds ratio | | |
|---|---|---|---|---|
| | *B(SE)* | Lower | Odds Ratio | Upper |
| Included | | | | |
| Constant | 0.08<br>(0.24) | | | |
| Eyes | 0.42<br>(0.34) | 0.78 | 1.52 | 3.00 |
| Camera | 0.10<br>(0.34) | 0.56 | 1.10 | 2.17 |

Note. $R^2$= .01 (Hosmer-Lemeshow), .01 (Cox-Snell), .01 (Nagelkerke), *Model* $\chi^2(2)$ 1.61, $p < .44$

Figure 4.4. shows the sample divided by condition and the percentage that chose to destroy at each possible destruction frequency (0-10). Most participants did not act maliciously, choosing not to destroy regardless of condition. The next most popular strategy involved choosing to destroy 50% of the time. Our data demonstrates that that some 116 of our 206 participants made at least one malicious decision when there was no benefit to them in doing so, but that there was no effect of our surveillance variables on such behaviour.

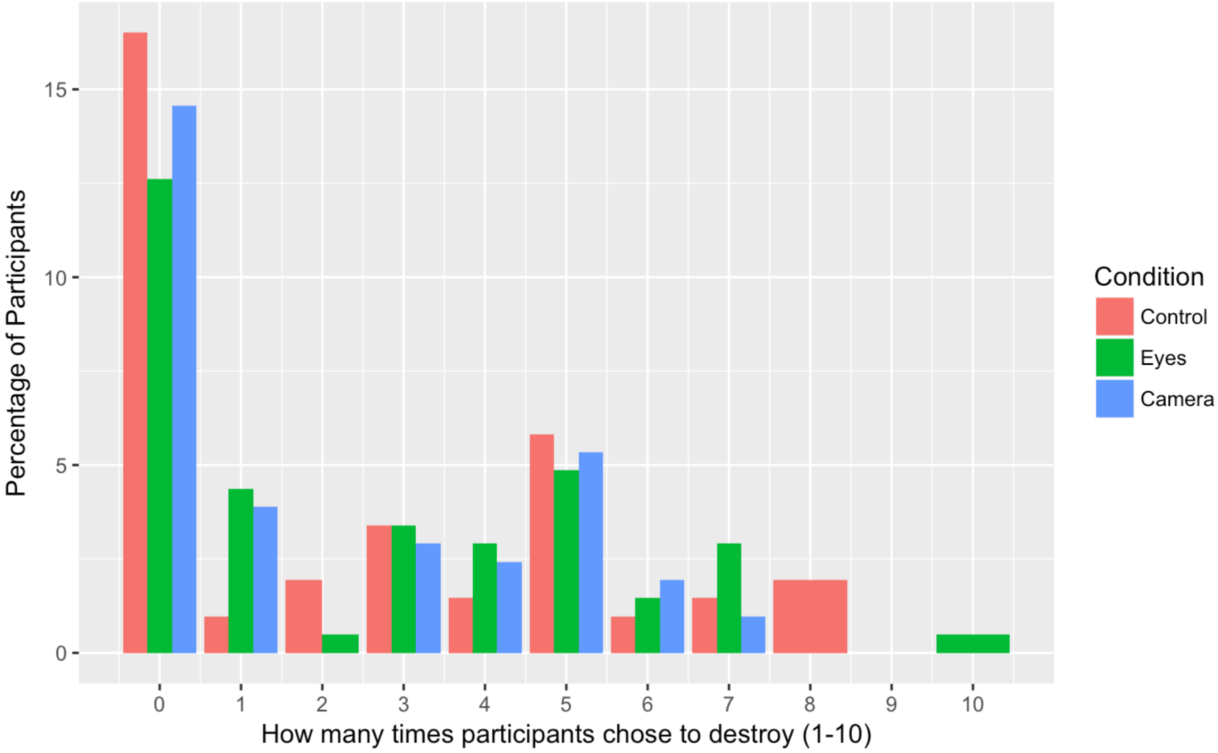


*Figure 4.4. Study 2. Distribution of Destruction Frequency by Participant.*

Figure 4.5. shows that there was a significant positive correlation between how frequently participants expected others to destroy their allocation and how often they chose to destroy others allocation themselves, ($r(206) = .53$, $p = <0.0001$). The correlation between expected destruction and actual destruction was as expected – the more participants expected others to destroy their allocation the more frequently they destroyed. Beyond the simple correlation, it is notable that participants expected their own allocation be to destroyed more often than they chose to destroy others (*Mdn Expected: Control = 0.4, Eyes = 0.4, Camera = 0.30; Mdn Destroyed: Control = 0.1, Eyes = 0.1, Camera = 0.15),* perhaps suggesting destruction was motivated not by maliciousness but by inequality aversion.

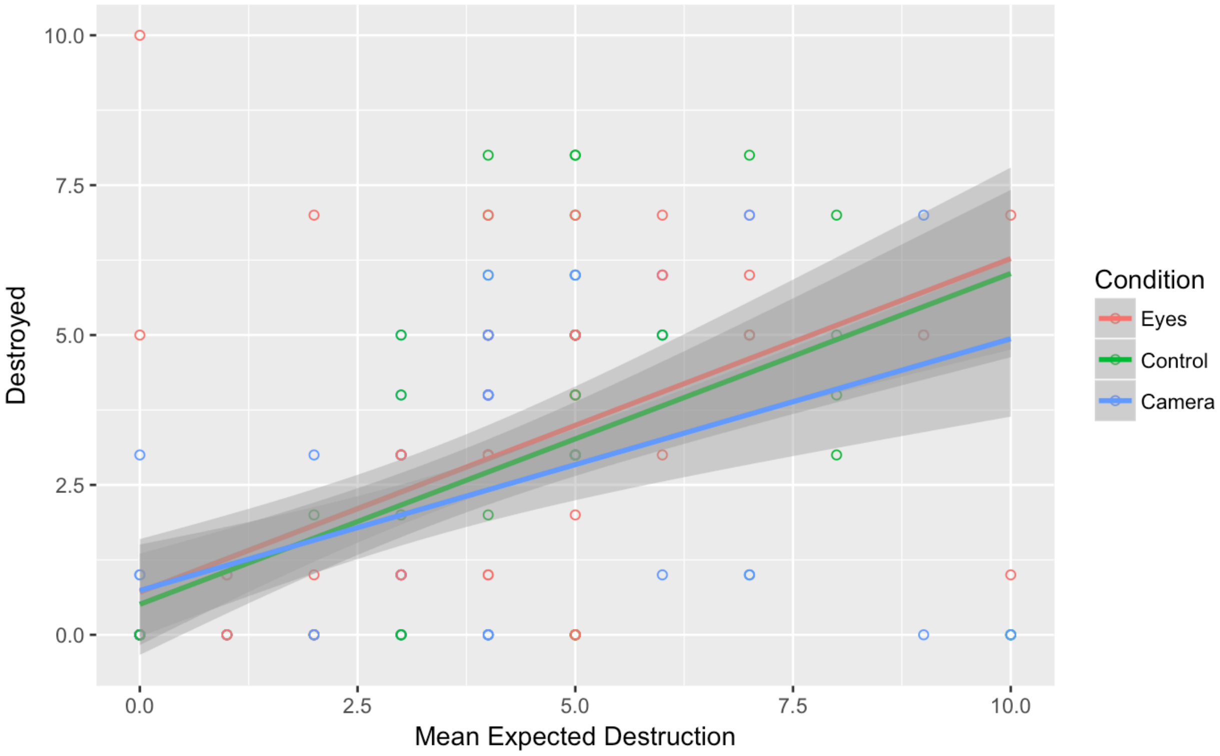


*Figure 4.5. Study 2. Mean Expected Destruction by Mean Destruction*

### 4.2.2.1. Public Self Consciousness and Surveillance Cues

Public self-consciousness in the Eyes (*N*=69, *Mdn*=11.00, *IQR* = 8.00, 15.00) and Camera (*N*=66, *Mdn* = 11.00, *IQR* = 7.00, 14.00) conditions, did not significantly differ from self-reported public-self-consciousness in the Control (*N* = 71, *Mdn* =10.00, *IQR* = 8.00, 13.50), *H(2) = 0.82, p* = .66.

### 4.2.2.2. Public Self-Consciousness as Moderator of Surveillance Cue Effects on Destruction Rates

A Poisson GLM examining whether public self-consciousness moderated the strength or direction of destruction rates found one significant interaction effect between the presence of the camera and destruction rates *b* = -0.64, *SE* = 0.02, *Z = 10.36, p* = .04 (see Table 4.5. & Figure 4.6. below). However, the overlapping confidence intervals (95% CI shown in grey for each condition in Figure 4.6.) give cause for caution.

Table 4.5. Study 2. Poisson GLM Examining Destruction rates in the JoD by Condition, Public Self-Consciousness

| | *B* | *SE B* | *Z* | *P* |
|---|---|---|---|---|
| Constant | 0.63 | 0.20 | 3.13 | <.001 *** |
| Condition: Eyes | 0.16 | 0.28 | 0.57 | .57 |
| Condition: Camera | -0.64 | 0.31 | -2.09 | .04* |
| Public Self-Consciousness | 0.02 | 0.02 | 1.23 | .22 |
| Condition Eyes: Public Self-Consciousness | -0.01 | 0.02 | -0.36 | .72 |
| Condition Camera: Public Self- Consciousness | 0.05 | 0.02 | 1.84 | .07 . |

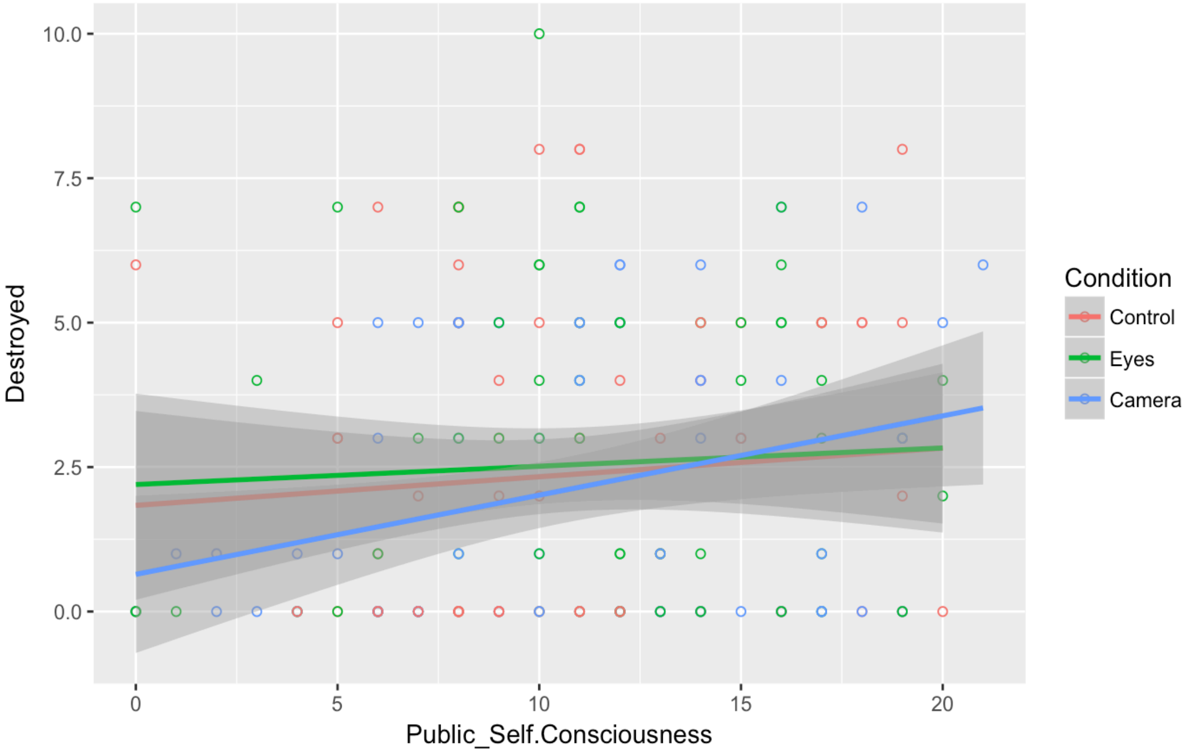


*Figure 4.6. Study 2. The influence of Public Self-Consciousness on Destruction Rates in the Joy of Destruction Game.*

### 4.2.2.3. Independence-Interdependence Scores as Moderator of Surveillance Cue Effects on Destruction Rates

The extent to which individuals' own identity was independent of, or overlapped with, their group identity was scored on the Singelis Self-Construal (SSC) scale from independent to interdependent. Our Poisson GLM examined the effect of self-construal scores and condition on rates of destruction. There was a significant interaction between condition, destruction rates and participants' score on the continuous independent-interdependent self-construal scale *b* = -0.45, *SE* = 0.13, *Z = 3.63, p* = < .001 (See Table 4.6 & Figure 4.7).

Table 4.6. Study 2. Poisson GLM Examining Destruction rates in the JoD by Condition, Independence-Interdependent Self-Construal

| | *B* | *SE B* | *Z* | *P* |
|---|---|---|---|---|
| Constant | 0.83 | 0.08 | 10.29 | <.001 *** |
| Condition: Eyes | 0.05 | 0.11 | 0.46 | .64 |
| Condition: Camera | -0.09 | 0.12 | -0.77 | .44 |
| Independence-Interdependence | 0.18 | 0.08 | 2.36 | .02* |
| Condition Eyes: Independence-Interdependence | 0.08 | 0.12 | 0.69 | .49 |
| Condition Camera: Independence-Interdependence | -0.45 | 0.13 | -3.63 | <.001 *** |

*Figure 4.7. Study 2. The influence of Independent-Interdependent Self-Construal on Destruction Rates in the Joy of Destruction Game.*

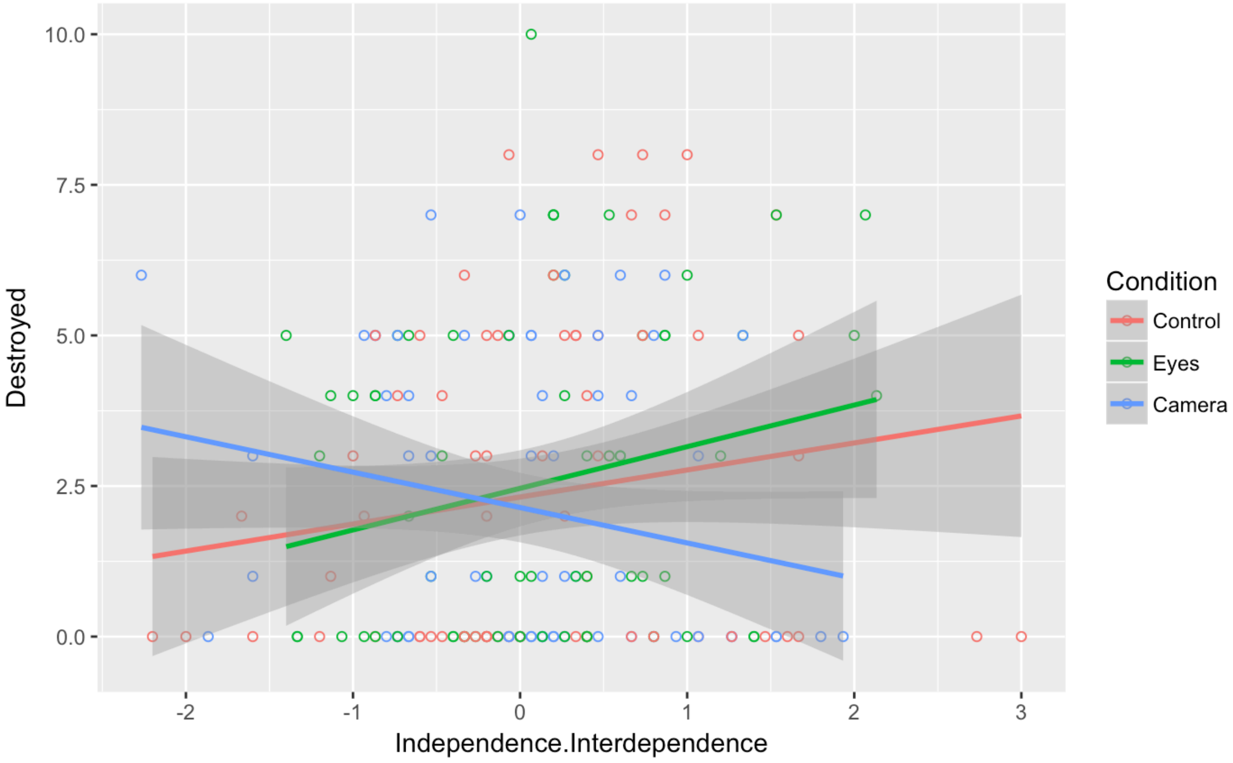


Figure 4.7. shows that in the Control and Eyes conditions increasing interdependence correlated with increasing destruction, whereas in the camera increasing interdependence scores correlated with decreasing destruction.

### 4.2.2.4. Social Value Orientation as Moderator of Surveillance Cue Effects on Destruction Rates

A Poisson regression found a significant interaction between SVO Scores, the eyes condition and destruction rates, $b = -0.04$, $Z = -2.48$, $SE = 0.02$, $p = .01$ (see Table 4.7 & Figure 4.8).

Table 4.7. Study 2. Poisson GLM Examining Destruction rates in the JoD by Condition, SVO Score

| | *B* | *SE B* | *Z* | *P* |
|---|---|---|---|---|
| Constant | 1.82 | 0.49 | 3.71 | <.001*** |
| Condition: Eyes | 1.76 | 0.67 | 2.62 | <.01** |
| Condition: Camera | -0.94 | 0.67 | -1.40 | .16 |
| SVO Score | -0.02 | 0.01 | -1.96 | .05* |
| Condition Eyes: SVO Score | -0.04 | 0.02 | -2.48 | .01* |
| Condition Camera: SVO Score | 0.02 | 0.02 | 1.22 | .22 |

*Figure 4.8. Study 2. The influence of Social Value Orientation on Destruction Rates in the Joy of Destruction Game.*

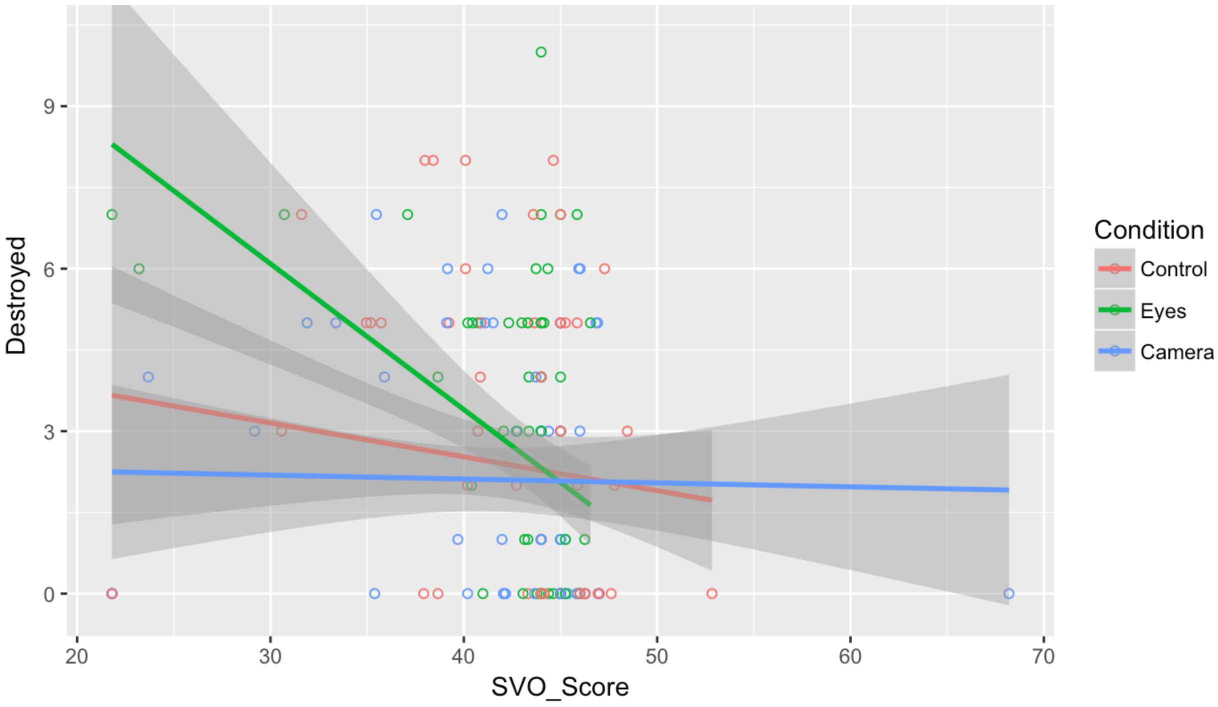


Figure 4.8. shows that in the eyes condition increasing prosociality scores correlated with a sharp decrease in destruction, whereas in the camera and control conditions the decline in destruction rates with increasing prosociality was much less steep. Removing outliers with

SVO scores more than 3 standard deviations above or below the mean made no difference to the relationship.

### 4.2.2.5. Religiosity as Moderator of Surveillance Cue Effects on Destruction Rates

We predicted there would be no difference in destruction rates by condition for highly religious participants because they would feel 'watched' by God in the control condition. In contrast low religious participants would destroy more frequently in the control than in the eyes and camera conditions, because the absence of a surveillance cue in the control condition would allow a greater feeling of anonymity and therefore less inhibition of destructive behaviour.

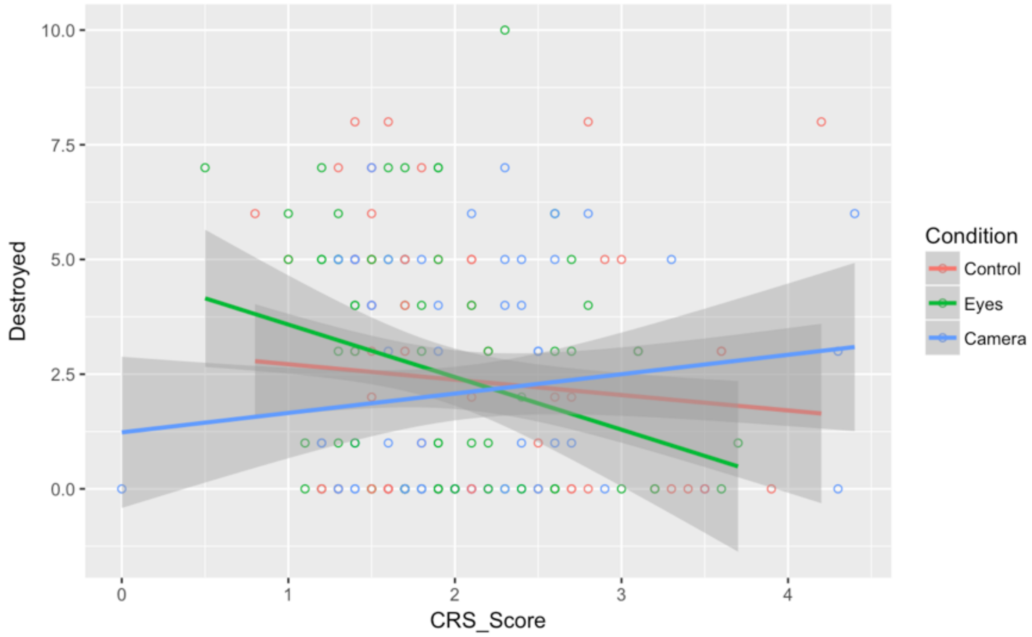
A Poisson regression found that destruction rates were significantly predicted by a combination of the presence of Eyes and CRS scores, *b* = -0.35, *SE* = 0.17, *Z* = -2.02, *p* = .04 and also by a combination of the presence of a Camera condition and CRS scores, albeit in the opposite direction *b* = 0.34, *SE* = 0.15, *Z* = 2.26, *p* = .02, see Table 4.8 & Figure 4.9.

Table 4.8. Study 2. Poisson GLM Examining Destruction rates in the JoD by Condition, Religiosity

| | *B* | *SE B* | *Z* | *P* |
|---|---|---|---|---|
| Constant | 1.16 | 0.23 | 5.08 | <0.001*** |
| Condition: Eyes | 0.67 | 0.33 | 2.03 | 0.04* |
| Condition: Camera | -0.81 | 0.33 | -2.47 | 0.01* |
| Centrality of Religion Score (CRS) | -0.15 | 0.11 | -1.37 | 0.17 |
| Condition Eyes: CRS | -0.35 | 0.17 | -2.02 | 0.04* |
| Condition Camera: CRS | 0.34 | 0.15 | 2.26 | 0.02* |

Figure 4.9 shows that in the control condition increasing religiosity correlated with a slight decrease in destruction rates. In the eyes condition the decrease in destruction rates with increasing religiosity was greater than in the control. Whereas in the camera condition the correlation was in the opposite direction - destruction increased as religiosity increased.

*Figure 4.9. Study 2. The influence of Centrality of Religion on Destruction Rates in the Joy of Destruction Game.*



### 4.2.3. Study 2 Discussion.

The absence of any main effect on destruction fails to replicate the findings of Baillon et al. (2013) who found that watching eyes predicted lower destruction rates in a watching eyes condition than in a control. Additionally, the absence of any main effect of surveillance cues on public self-consciousness failed to replicate Pfattheicher and Keller (2015) finding that the presence of watching eyes made people feel more self-conscious and thus, as they inferred, more watched.

The significant interaction effects we found were uninformative. The significant effect of public self-awareness and camera condition on destruction was in the opposite direction to that we had predicted. We predicted that as public self-consciousness increased participants would destroy less in the eyes and camera condition where they would be cued to manage their reputation. In contrast we predicted that in the control there would be no correlation between public self-consciousness and destruction since all would feel equally anonymous – low public self-consciousness individuals wouldn't be managing the impression they make on others because they don't anyway, high public self-consciousness individuals wouldn't be managing the impression they make on others because there was no cue to trigger this disposition. Figure 4.6. shows that in fact destruction increased (rather the reduced) as public self-consciousness rose in all conditions including the control. Indeed, the results are the opposite of what we would predict – those most likely to adapt their behaviour to protect their reputation when watched destroyed more with a camera or eye cues present (and more in the control), not less.

The significant interaction effect between self-construal score and condition on destruction rates was as predicted in the camera condition but not in the eyes and control conditions. In the camera condition the more interdependent an individual's identity was with their group, and consequently the more responsible they felt for managing the collective reputation, the less they destroyed. However, in the control and eyes condition destruction increased as interdependence increased. We had expected a similar correlation in the eyes condition to that in the camera – greater interdependence predicting lower destruction rates – and no increase, or only a small increase in destruction rates with rising interdependence in the control.

Only one of the interaction effects found between SVO scores and condition on destruction rates was as expected. The significant interaction between SVO Score and condition on destruction rates was driven by the sharp reduction in destruction rates with increasing prosociality in the camera condition. There were smaller correlations between rising prosociality and lower destruction in the control and eyes condition. It *was* expected that prosociality would predict lower destruction rates in the control condition, where decisions to destroy or not were expected to be a function of intrinsic motivation only. However, in contrast, in the eyes and camera condition it was expected that low prosocials would destroy less as low prosocials subconsciously adapted their behaviour and destroyed less in order to manage their reputation. In fact, the correlation between increasing prosociality and lower destruction rates was strongest in the camera condition.

We predicted that in the control condition, increasing religiosity would predict lower destruction rates since the highly religious might always feel ‘watched by god’ whereas lower religious participants would not feel watched. Consequently, when a surveillance cue was present and all participants were cued to feel watched – by God and/or eyes or a camera - there would be no correlation between religiosity and destruction rates. In fact, in the camera condition increasing religiosity predicted higher destruction. In the eyes conditions this effect, though small, was reversed: increasing religiosity predicted lower destruction where we had expected no effect. In the control condition the correlation with increasing religiosity predicting lower destruction was present, but was not as strong as in the eyes condition, again contradicting our prediction.

Taken together, noting the overlapping confidence intervals and effects in the opposite direction to our hypotheses, the statistically significant differences required replication to be interpreted meaningfully.

## 4.3. Study 3

Study 3 was an online study which examined the effect of watching eyes or a camera on participants malicious behaviour in the Joy of Destruction task, aiming to replicate the findings of Study 2.

### 4.3.1. Methods

#### 4.3.1.1. Participants

Participants were all over the age of 18 and recruited online via Prolific Academic (Palan & Schitter, 2018). Our sample in Study 3 was 47% male (*N*=91), 53% female (*N* = 102), 0.5% (*N* = 1) other, or did not report their gender. The mean age of participants was 30 years old. 68% were Caucasian, White or White European. A further 16% reported a European nationality (e.g. British, Greek, Polish) without indicating an ethnicity suggesting as many as 84% might have been White European. Just 4% reported an Asian ethnicity or nationality (e.g. Chinese, Indian), 3% reported various mixed ethnicities, 3% identified as Black British, 3% as Latin, Hispanic or Mexican. The remainder did not respond to the ethnicity question.

#### 4.3.1.2. Design

In study 3 we sought to extend the findings from Study 2 in an online environment to confirm that online testing did not produce different results to those found under laboratory conditions.

Study 3's design was identical to Study 2 in all possible respects. Any changes were the unavoidable consequence of moving a laboratory study online (see subparagraph 4.3: *Stimuli*). As a direct replication of Study 2, it used the same software. Participants first completed the JoD task, followed by the individual difference measures.

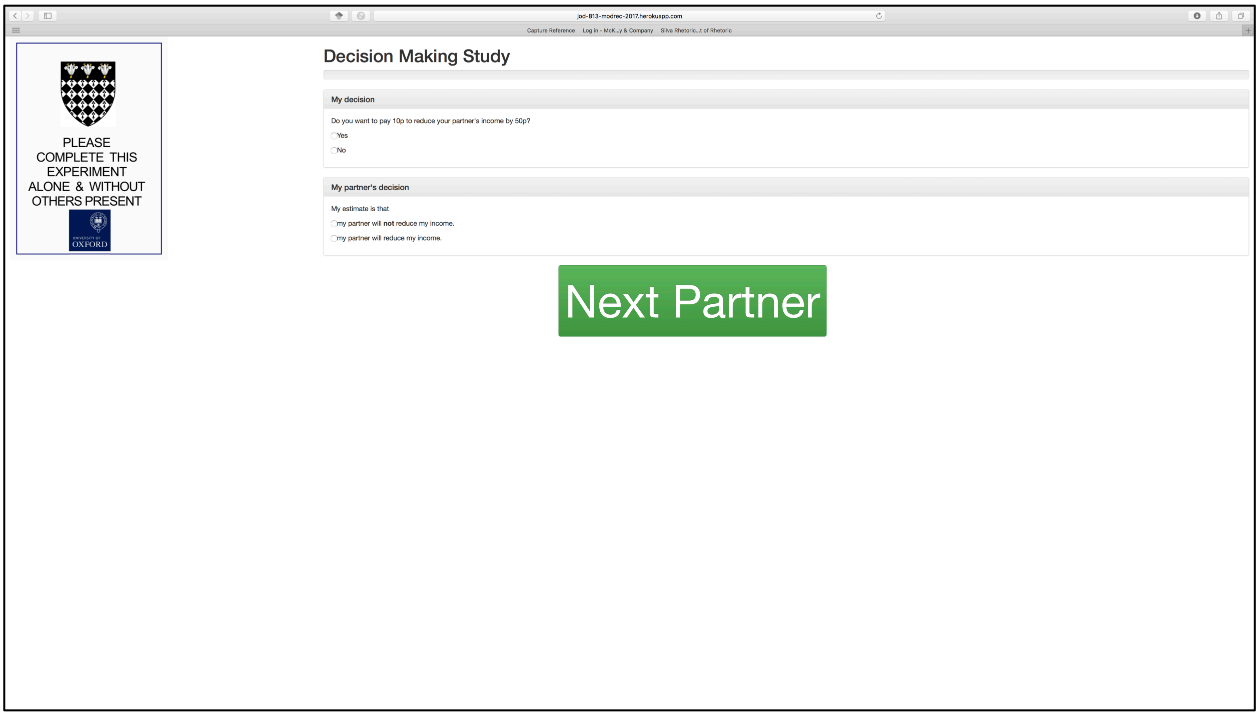
#### 4.3.1.3. Materials

All materials were identical to Study 2.

#### 4.3.1.4. Stimuli

On screen stimuli were identical to Study 2. However, being an online experiment the additional eye cues and control images on the posters on the walls of the booth in the laboratory conditions study were not present. The text accompanying the image was changed from ‘please do not eat or drink in the cubicle’ to ‘please complete this experiment alone and without others present’ to reflect the fact that it was being run online (see Figure 4.10 below). Additionally, the Logitech camera used in the laboratory conditions study was replaced by the integrated webcam on the devices of our online participants. We could not control where participants completed the experiment hence the on-screen request that they do so alone and without others present.

*Figure 4.10. Screenshots as seen by participants in studies 2, 3, & 4. Images displayed show the control (top) and eye cues (bottom) condition. The control image screen was also used for the camera condition, but with the participant's webcam switched on too.*


jod-813-modrec-2017.herokuapp.com
Capture Reference
Log in - McK...y & Company
Silva Rhetoric...t of Rhetoric
UNIVERSITY OF OXFORD
Next Partner

PLEASE
COMPLETE THIS
EXPERIMENT
ALONE & WITHOUT
OTHERS PRESENT

Decision Making Study

**My decision**

Do you want to pay 10p to reduce your partner's income by 50p?

○ Yes

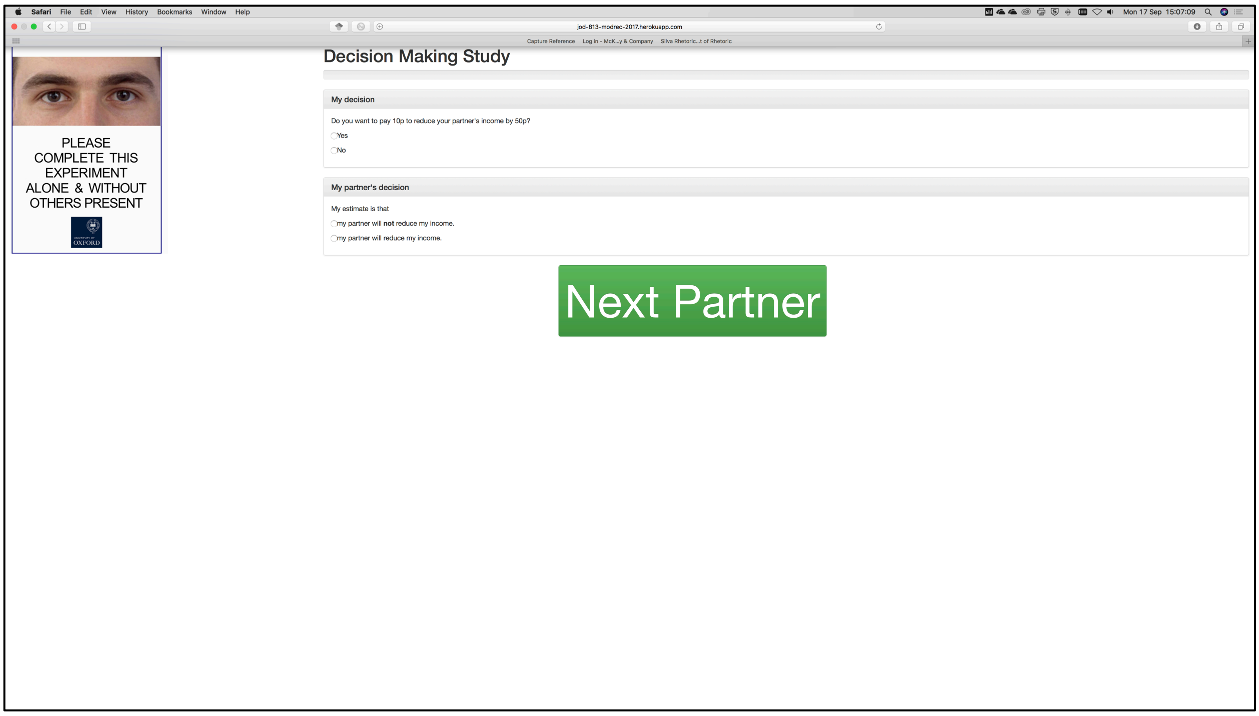
○ No

**My partner's decision**

My estimate is that

○ my partner will **not** reduce my income.

○ my partner will reduce my income.


Safari
File
Edit
View
History
Bookmarks
Window
Help
Mon 17 Sep 15:07:09
jod-813-modrec-2017.herokuapp.com
Capture Reference
Log in - McK...y & Company
Silva Rhetoric...t of Rhetoric
UNIVERSITY OF OXFORD
Next Partner

PLEASE
COMPLETE THIS
EXPERIMENT
ALONE & WITHOUT
OTHERS PRESENT

Decision Making Study

**My decision**

Do you want to pay 10p to reduce your partner's income by 50p?

○ Yes

○ No

**My partner's decision**

My estimate is that

○ my partner will **not** reduce my income.

○ my partner will reduce my income.

#### 4.3.1.5. Procedure

As with study 2, all data was recorded by the software version of the JoD (Dear & Ellen, 2018).

#### 4.3.1.6. Data Plan

Assuming the data would be non-parametric again, statistical methods were adopted from Study 2 employing principally Kruskal-Wallace tests, Logistic Regression, and Poisson GLMs.

### 4.3.2. Study 3 Results

Participant destruction rates did not differ significantly between our 2 experiments, $H(1) = 0.44$, $p = 0.50$. Table 4.9. compares results in Study 3 with results in Study 2. As can be seen, the only significant differences in responses to our questionnaire measures was in the reversal of the direction of self-construal scores *(*Study 2 $M_{\mathit{Independence-Interdependence(I-I)}} = 0.1$, Study 3 $M_{I\text{-}I} = -0.16$*)*.

Table 4.9. Summary Statistics Comparing Experiment 3 and Experiment 2

| | | Study 2 (Laboratory Study) | | Study 3 (Online Replication) | | Test Statistic | *P* | Test |
|---|---|---|---|---|---|---|---|---|
| *N* | | 206 | | 194 | | | | |
| Male/Female/Unknown | | 137/64/5 | | 91/102/1 | | | | |
| Destruction (0-1) | median [*IQR*] | 0.10 | [0.00, 0.50] | 0.20 | [0.00, 0.50] | *H* = 0.44 | .51 | K-W |
| Expected Destruction | median [*IQR*] | 0.40 | [0.00, 0.50] | 0.40 | [0.00, 0.50] | *H* = 0.19 | .67 | K-W |
| Public Self-Consciousness | median [*IQR*] | 11.00 | [6.00, 20.00] | 10.00 | [6.00, 15.00] | *H* = 1.35 | .25 | K-W |
| Independence-Interdependence | mean (*sd*) | 0.10 | (0.88) | -0.16 | (0.88) | *F* = 8.99 | <.01** | ANOVA |
| SVO Score | median [*IQR*] | 44.00 | [42.05, 45.00] | 44.00 | [40.46, 45.00] | *H* = 2.97 | .08 | K-W |
| SVO Cat | Prosocial (%) | 196 | (95.1) | 186 | (95.9) | | | |
| CRS Score | median [*IQR*] | 1.80 | [1.50, 2.40] | 1.90 | [1.40, 2.80] | *H* = 1.05 | .31 | K-W |

Values presented refer to the median, inter-quartile range, p-value and significance test associated with each variable.

In Study 3 destruction rates were again significantly non-normal, both in the aggregate, $W = 0.85$, $p < .001$ and by condition, Control $W = 0.83$, $p < .001$, Eyes $W = 0.87$, $p < .001$, Camera, $W = 0.84$, $p < .001$. Levene's test indicated that the variances in destruction rates by condition were not significantly differently, $F(2, 191) = 2.69$, $p = .07$.

There was no significant main effect of condition on median destruction rates in Study 3, $H(2) = 0.49$, $p =$ *0.78.* Table 4.10 shows the results of Kruskal Wallace tests for the non-parametrically distributed variables and a one-way ANOVA for the normally distributed Independence-Interdependence scores. Neither destruction rates, public self-consciousness scores nor our other questionnaire results differed significantly by condition.

Table 4.10. Study 3 Results Stratified by Condition

| | | Control | | Eyes | | Camera | | Test Statistic | *P* | Test |
|---|---|---|---|---|---|---|---|---|---|---|
| *N* | | 75 | | 64 | | 55 | | | | |
| Male/Female/Unknown | | 33/41/1 | | 27/37/0 | | 31/24/0 | | | | |
| Destruction (0-1) | median [*IQR*] | 0.20 | [0.00, 0.50] | 0.20 | [0.00, 0.50] | 0.20 | [0.00, 0.40] | $H = 0.95$ | 0.62 | K-W |
| Expected Destruction | median [*IQR*] | 0.30 | [0.00, 0.50] | 0.40 | [0.15, 0.60] | 0.40 | [0.12, 0.50] | $H = 3.18$ | 0.20 | K-W |
| Public Self-Consciousness | median [*IQR*] | 8.00 | [5.00, 13.00] | 7.00 | [4.00, 11.00] | 9.00 | [4.50, 17.50] | $H = 3.46$ | 0.18 | K-W |
| Independence-Interdependence | mean (*sd*) | -0.21 | (0.81) | -0.10 | (1.00) | -0.17 | (0.84) | $F = 0.25$ | 0.78 | ANOVA |
| SVO Score | median [*IQR*] | 44.00 | [39.18, 45.00] | 44.00 | [41.73, 45.06] | 44.00 | [40.65, 44.87] | $H = 1.59$ | 0.45 | K-W |
| SVO Cat | Prosocial (%) | 72 | (96.0) | 61 | (95.3) | 53 | (96.4) | | 0.95 | |
| CRS Score | median [*IQR*] | 2.10 | [1.40, 2.95] | 1.95 | [1.50, 2.90] | 1.80 | [1.35, 2.50] | $H = 3.03$ | 0.22 | K-W |

Values presented refer to the median, inter-quartile range, p-value and significance test associated with each variable.

Logistic regression showed that the Surveillance cues did not predict destruction better than chance, Model $\chi^2(2)$ 1.12, *p* =.57. There was therefore no significant difference in the probability of destroying or not by condition (Eyes *z* = -0.97, *p* = 0.34; Camera *z* = 0.04, *p* = .97). The results of the logistic regression are in Table 11.

Table 4.11. Study 3. Logistic Regression Condition vs Destroyed or Not

| | | 95% CI for odds ratio | | |
|---|---|---|---|---|
| | *B(SE)* | Lower | Odds Ratio | Upper |
| Included | | | | |
| Constant | 0.24<br>(0.23) | | | |
| Eyes | 0.34<br>(0.35) | 0.71 | 1.40 | 2.80 |
| Camera | 0.015<br>(0.36) | 0.50 | 1.01 | 2.05 |

Note. $R^2$= .004 (Hosmer-Lemeshow), .01 (Cox-Snell), .01 (Nagelkerke), Model $\chi^2(2)$ 1.12, *p* =.57

### 4.3.2.1. Public Self Consciousness and Surveillance Cues

Public self-consciousness in the Eyes (*N*=71, *Mdn*=8.00, *IQR* = 4.00, 12.00) and Camera (*N*=62, *Mdn* = 9.00, *IQR* = 5.00, 17.00) conditions, didn't significantly differ from that reported in the Control (*N* = 77, *Mdn* =8.50, *IQR* = 5.00, 13.00), *H(2) = 3.46, p* = .18.

### 4.3.2.2. Public Self-Consciousness as Moderator of Surveillance Cue Effects on Destruction Rates

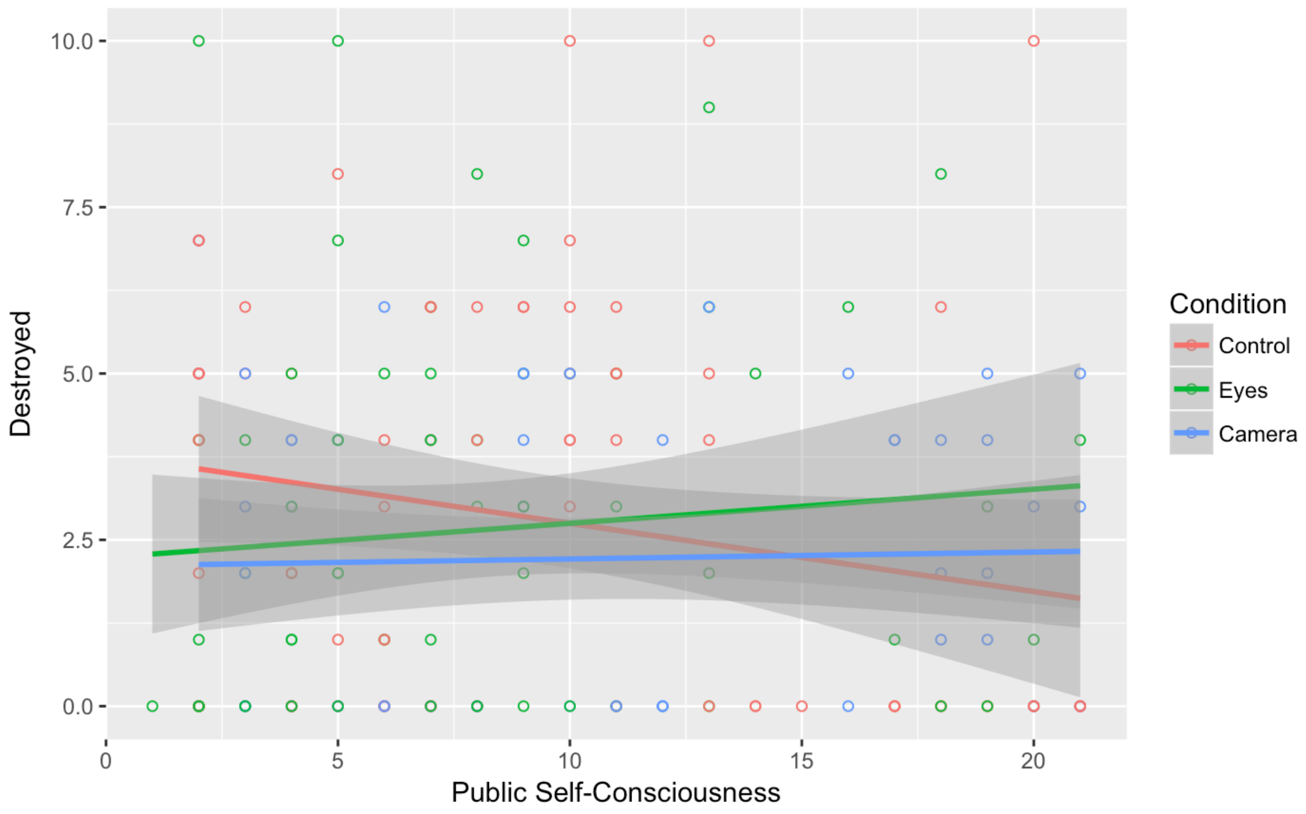
As with Study 2, we constructed a Poisson GLM examining interaction effects between condition, public self-consciousness and destruction rates. We found significant interactions effects between the Eyes condition and Public Self-Consciousness on destruction, *b* = 0.06, *SE* = 0.02, *p*

$< 0.01$, and the Camera Condition and public self-consciousness, $b = 0.04$, $SE = 0.02$, $p = .02$ (see Table 4.12 & Figure 4.11).

Table 4.12. Study 3. Poisson GLM Examining Destruction rates in the JoD by Condition, Public Self-Consciousness

| | *B* | *SE B* | *Z* | *P* |
|---|---|---|---|---|
| Constant | 1.38 | 0.13 | 10.98 | <0.001*** |
| Condition: Eyes | -0.56 | 0.19 | -2.93 | <0.01** |
| Condition: Camera | -0.63 | 0.21 | -2.95 | <0.01** |
| Public Self-Consciousness | -0.04 | 0.02 | 3.09 | <0.01** |
| Condition Eyes: Public Self-Consciousness | 0.06 | 0.02 | 3.09 | <0.01** |
| Condition Camera: Public Self- Consciousness | 0.04 | 0.02 | 2.35 | 0.02* |

*Figure 4.11. Study 3. The influence of Public Self-Consciousness on Destruction Rates in the Joy of Destruction Game.*



### 4.3.2.3. Interdependence Scores as Moderator of Surveillance Cue Effects on Destruction Rates

There was no significant effect of interdependence scores and surveillance cues on destruction rates, as can be seen our Poisson GLM in Table 4.13.

Table 4.13. Study 3. Poisson GLM Examining Destruction rates in the JoD by Condition, Independence-Interdependent Self-Construal

| | *B* | *SE B* | *Z* | *P* |
|---|---|---|---|---|
| Constant | 1.03 | 0.07 | 14.45 | <0.001*** |
| Condition: Eyes | -0.05 | 0.10 | -0.48 | 0.63 |
| Condition: Camera | -0.22 | 0.12 | -1.87 | 0.06. |
| Independence-Interdependence | 0.002 | 0.09 | 0.02 | 0.98 |
| Condition Eyes: Independence-Interdependence | 0.17 | 0.12 | 1.49 | 0.14 |
| Condition Camera: Independence-Interdependence | 0.13 | 0.14 | 0.93 | 0.35 |

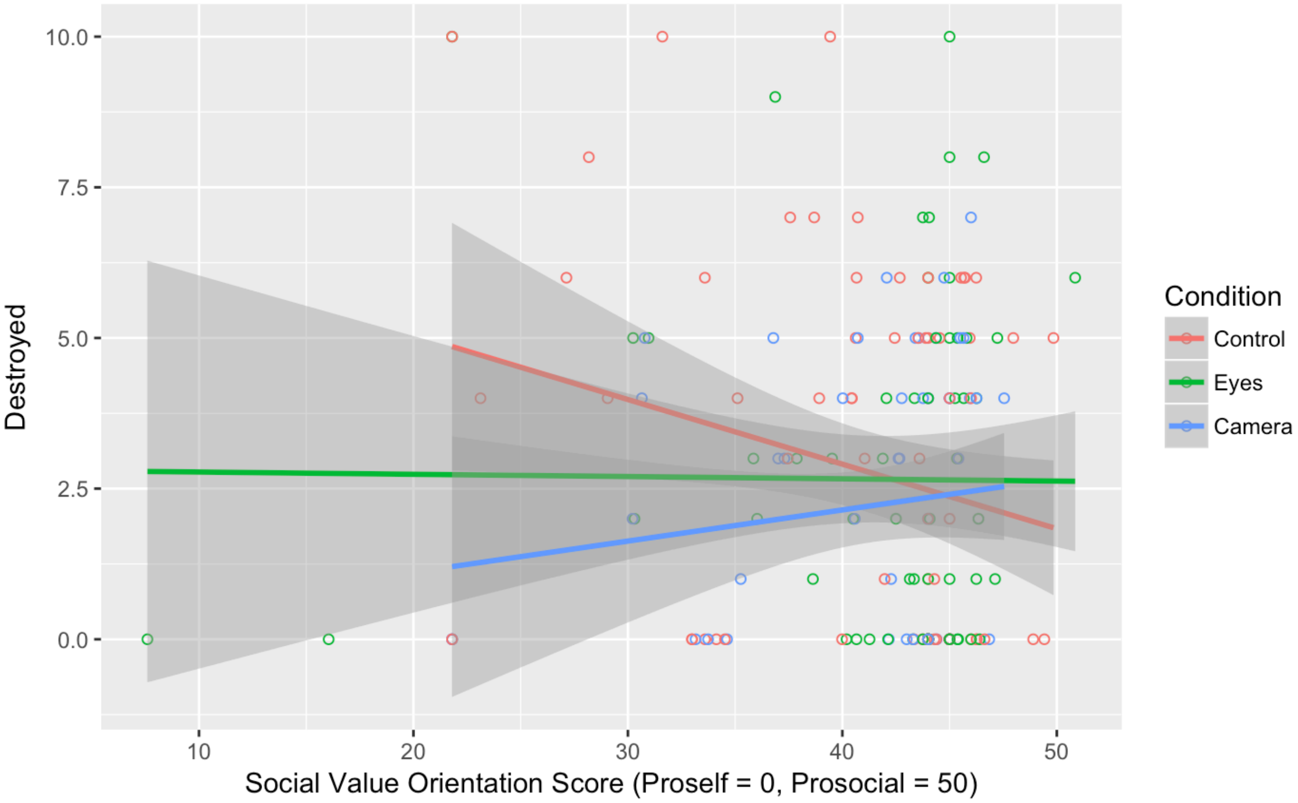
### 4.3.2.4. Social Value Orientation as Moderator of Surveillance Cue Effects on Destruction Rates

In Study 3 our Poisson GLM (Table 4.14) showed that there was a significant interaction effect between SVO Score and surveillance condition on destruction rates both in the eyes ($b$ = 0.03, $SE$ = 0.01, $p$ = .02) and camera conditions ($b$ = 0.06, $SE$ = 0.02, $p$ < 0.01). Figure 4.12 shows that there was no interaction effect between SVO Scores and the Eyes condition on destruction. It shows that there was, however, an SVO condition interaction effect with both camera and control conditions in predicting destruction: In the control condition higher SVO scores predicted decreased destruction, in the camera condition they predicted increased destruction.

Table 4.14. Study 3. Poisson GLM Examining Destruction rates in the JoD by Condition, SVO Score

| | *B* | *SE B* | *Z* | *P* |
|---|---|---|---|---|
| Constant | 2.34 | 0.36 | 6.51 | <0.001*** |
| Condition: Eyes | -1.31 | 0.57 | -2.28 | 0.02* |
| Condition: Camera | -2.67 | 0.86 | -3.11 | 0.002** |
| SVO Score | -0.03 | 0.01 | -3.64 | <0.001*** |
| Condition Eyes: SVO Score | 0.03 | 0.01 | 2.26 | 0.02* |
| Condition Camera: SVO Score | 0.06 | 0.02 | 2.92 | <0.01** |

*Figure 4.12. Study 3. The influence of Social Value Orientation on Destruction Rates in the Joy of Destruction Game.*



### 4.3.2.5. Religiosity as Moderator of Surveillance Cue Effects on Destruction Rates

In Study 3 there was no significant interaction effect between either surveillance cue and centrality of religion on destruction rates (Table 4.15).

Table 4.15. Study 2. Poisson GLM Examining Destruction rates in the JoD by Condition, Religiosity

| | *B* | *SE B* | *Z* | *P* |
|---|---|---|---|---|
| Constant | 1.16 | 0.17 | 6.68 | <0.001*** |
| Condition: Eyes | -0.46 | 0.25 | -1.82 | 0.07 . |
| Condition: Camera | -0.72 | 0.30 | -2.41 | 0.41 |
| Centrality of Religion Score (CRS) | -0.06 | 0.07 | -0.82 | 0.41 |
| Condition Eyes: CRS | 0.18 | 0.10 | 1.76 | 0.08 . |
| Condition Camera: CRS | 0.23 | 0.13 | 1.80 | 0.07 . |

## 4.3.3. Study 3 Discussion

In study 3 we sought to extend the findings from Study 2 in an online environment to confirm that online testing did not produce different results to those found under laboratory conditions. There were no significant differences in reported destruction rates nor in self-awareness scores between study 2 and 3 suggesting that the online study was able to replicate the offline results.

As seen in Figure 4.12, in Study 3 higher public self-consciousness predicted lower destruction rates in the camera condition, but higher destruction rates in the control condition, whereas there was no correlation in the eyes condition. The slope for the control condition was in the opposite direction to that we found in Study 2. The absence of any correlation between higher public self-conscious scores and lower destruction in the control also violates our hypothesis. The effect in the camera condition is in the direction predicted, with higher self-consciousness scores predicting lower destruction when a camera was present. But this is again in the opposite direction to what we found in Study 2. These conflicting results, and the wide confidence intervals, discourage further extrapolation from our findings.

Similarly, the interactions between surveillance cues and social value orientation were not consistent between studies 2 and 3. Indeed the direction of the slopes for both the control condition and the camera condition were reversed between study 2 and 3, while the confidence intervals again overlapped.

The different findings in the exploratory analyses between Study 2 & 3 are most likely the result of false positives. We explore this in Study 4 to see if Study 3's effects replicate in a second online experiment.

Both Study 2 & 3 might have found no effect of surveillance cues due to the psychological distance between participants induced by the 'hidden condition'. Noting this limitation ourselves and responding to feedback from one of the originators of the Joy of Destruction task (Hermann, personal communication, February 5, 2018) we designed a third study to test whether eye cues or a camera would influence decisions when the role of chance was removed – thus participants would know that any reduction in their allocation was not bad luck but the result of the decision by another participant. Similarly, participants would know that if they paid to reduce another's income, their victim could not write it of as chance and would know that someone else had paid to reduce their income.

## 4.4. Study 4

Study 4 was an online study which examined the effect of watching eyes or a camera on participants malicious behaviour in the Joy of Destruction task, this time removing the role of chance in calculating participant's losses.

### 4.4.1. Methods

#### 4.4.1.1. Participants

As with Study 3 participants were all over the age of 18 and recruited via Prolific Academic (Palan & Schitter, 2018).

Our sample in Study 4 was 54% male (N=104), 45% female (N = 87), 1 participant did not supply gender. The mean age of participants was 30 years old. 70% of participants could be classified as Caucasian, White or White European, and a further 17% gave a European nationality (e.g. Greek, Polish) without an ethnicity – thus as much as 87% may have been white European. Just 5% reported an Asian ethnicity or nationality (e.g. Chinese, Indian), 4% reported various mixed ethnicities, 2% identified as Black British. The remainder was 1% Hispanic, 1% Jewish.

#### 4.4.1.2. Design

Study 4 differed from Study 2 & 3 only by removing the role of chance in determining participant outcomes. Stimuli, data recording and statistical methods were identical with study 3.

#### 4.4.1.3. Materials

All materials were identical to Study 2 & 3.

#### 4.4.1.4. Stimuli

Stimuli were identical to Study 3.

#### 4.4.1.5. Procedure

As with study 2 & 3, all data was recorded by the software version of the JoD (Dear & Ellen, 2018).

#### 4.4.1.6. Data Plan

Assuming the data would be non-parametric again, statistical methods were adopted from Study 2 employing principally Kruskal-Wallace tests, Logistic Regression, and Poisson GLMs.

### 4.4.2. Results

In Study 4, destruction rates were significantly non-normal, both in the aggregate, $W = 0.81$, $p < .001$ and by condition, Control $W = 0.83, p < .001$, Eyes $W = 0.79, p < .001$ , Camera, $W = 0.80$, $p < .001$. Levene's test indicated that variances in destruction rate by condition were not significantly differently, $F(2, 189) = 1.35, p = .26$. Table 4.16 provides summary statistics for Study 4.

Destruction rates did not differ significantly by condition across the three Joy of Destruction studies reported in this chapter $H(2) = 2.21, p = 0.33$.

Table 4.16. Study 4: JoD Results Stratified by Condition

| | | Control | | Eyes | | Camera | | Test Statistic | *P* | Test |
|---|---|---|---|---|---|---|---|---|---|---|
| *N* | | 70 | | 66 | | 56 | | | | |
| Male/Female/Unknown | | 34/35/1 | | 38/28/0 | | 32/24/0 | | | | |
| Destruction (0-1) | median [*IQR*] | 0.20 | [0.00, 0.50] | 0.05 | [0.00, 0.40] | 0.05 | [0.00, 0.40] | *H* = 2.21 | 0.33 | K-W |
| Expected Destruction | median [*IQR*] | 0.30 | [0.10, 0.50] | 0.30 | [0.00, 0.47] | 0.30 | [0.00, 0.50] | *H* = 2.36 | 0.31 | K-W |
| Public Self-Consciousness | median [*IQR*] | 13.00 | [9.25, 16.00] | 12.50 | [8.00, 17.00] | 13.00 | [10.00 17.00] | *H* = 0.47 | 0.79 | K-W |
| Independence-Interdependence | mean (*sd*) | -0.10 | (1.04) | -0.26 | (0.89) | -0.47 | (0.81) | *F* = 2.46 | 0.09 | |
| SVO Score | median [IQR] | 30.10 | [22.93, 36.12] | 26.44 | [12.66, 34.88] | 28.95 | [17.03, 34.88] | *H* = 3.91 | 0.14 | K-W |
| SVO Cat | Prosocial, *N* (%) | 52 | (74.3) | 41 | (62.1) | 39 | (69.6) | | | |
| CRS Score | median [IQR] | 1.85 | [1.50, 2.40] | 1.80 | [1.40, 2.38] | 2.00 | [1.58, 2.90] | *H* = 2.74 | 0.25 | K-W |

Values presented refer to the median, inter-quartile range, p-value and significance test associated with each variable.

Logistic regression showed that our model was no better at predicting destruction rates than chance, Model $\chi^2(2)$ 0.91, $p = .63$. Consequently, there was no significant difference in the probability of destroying or not by condition (Eyes $z = -0.83$, $p = .40$; Camera $z = -0.80$, $p = .43$). The results of the logistic regression are in Table 4.17.

Table 4.17. Study 4. Logistic Regression Condition vs Destroyed or Not

| | | 95% CI for odds ratio | | |
|---|---|---|---|---|
| | B(SE) | Lower | Odds Ratio | Upper |
| Included | | | | |
| Constant | 0.29 | | | |
| | (0.24) | | | |
| Eyes | -0.29 | 0.38 | 0.75 | 1.47 |
| | (0.34) | | | |
| Camera | -0.29 | 0.37 | 0.75 | 1.52 |
| | (0.36) | | | |

Note. $R^2$= .003 (Hosmer-Lemeshow), .01(Cox-Snell), .01 (Nagelkerke), Model $\chi^2(2)$ 0.91, $p = .63$

In contrast to previous studies (Abbink & Hermann, 2009; Abbink & Sadrieh, 2009) removing the role of chance in order that participants knew that their ‘partners’ would know that any destruction they suffered was the result of human agency made no significant difference to median destruction rates (Hidden vs Not Hidden, $H(1) = .90$, $p =$ *0.34*.

Median participant destruction rates did not differ significantly between our 3 experiments, $H(2) = 1.35$, $p = 0.51$.

#### 4.4.2.1. Public Self Consciousness and Surveillance Cues

Public self-consciousness in the Eyes (*N*=66, *Mdn*=12.50, *IQR* = 8.00, 17.00) and Camera (*N*=56, *Mdn* = 10.50, *IQR* = 6.00, 14.00) conditions, did not significantly differ from self-reported public-self-consciousness in the Control (*N* = 70, *Mdn* =13.00, *IQR* = 9.25, 16.00), *H(2) = 0.47, p* = .79.

#### 4.4.2.2. Public Self Consciousness as Moderator of Surveillance Cue Effects on Destruction Rates

As with Study 2 & 3, we constructed Poisson GLM's examining interaction effects between condition, public self-consciousness and destruction rates. In Study 4 there were no significant interaction effects between condition, public self-consciousness scores and destruction rates (see Table 4.18).

Table 4.18. Study 4. Poisson GLM Examining Destruction rates in the JoD by Condition, Public Self-Consciousness

| | *B* | *SE B* | *Z* | *P* |
|---|---|---|---|---|
| Constant | 0.72 | 0.23 | 3.21 | 0.001** |
| Condition: Eyes | 0.08 | 0.31 | 0.25 | 0.80 |
| Condition: Camera | -0.06 | 0.38 | -0.17 | 0.86 |
| Public Self-Consciousness | 0.02 | 0.02 | 1.25 | 0.21 |
| Condition Eyes: Public Self-Consciousness | -0.02 | 0.03 | -0.78 | 0.44 |
| Condition Camera: Public Self- Consciousness | -0.02 | 0.03 | -0.78 | 0.44 |

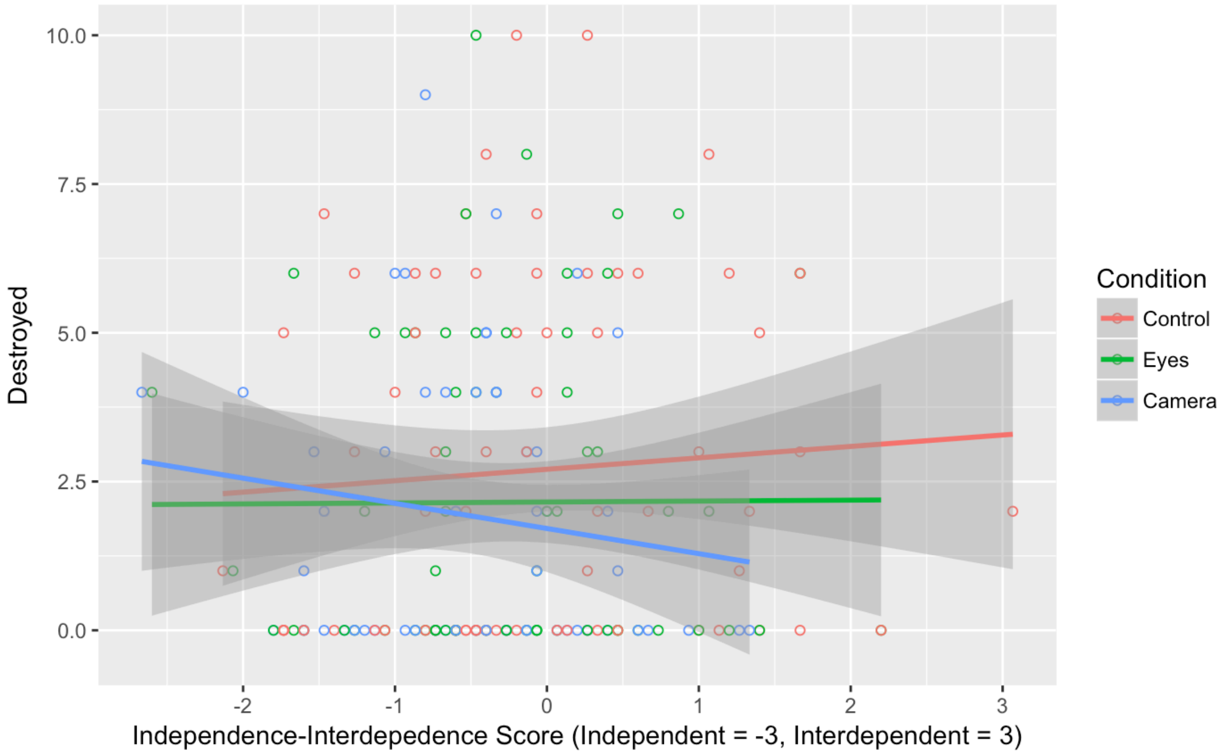
### 4.4.2.3. Interdependence Scores as Moderator of Surveillance Cue Effects on Destruction Rates

In Study 4 there was no significant interaction effect between the eyes condition, interdependence scores and destruction rates. However, the camera and interdependence score interaction did have a significant effect on destruction rates, *b* = -0.29, *SE* = 0.14, *p* = .03 (see Table 4.19.). Figure 4.13. shows that highly independent participants destroyed less in the camera condition than in the control, whereas the more interdependent participants destroyed at similar rates irrespective of condition.

Table 4.19. Study 3. Poisson GLM Examining Destruction rates in the JoD by Condition, Independent-Interdependent Self-Construal

| | *B* | *SE B* | *Z* | *P* |
|---|---|---|---|---|
| Constant | 0.99 | 0.07 | 13.61 | <0.001*** |
| Condition: Eyes | -0.22 | 0.11 | -1.97 | 0.05* |
| Condition: Camera | -0.46 | 0.14 | -3.28 | 0.001** |
| Independence-Interdependence | 0.07 | 0.07 | 1.01 | 0.31 |
| Condition Eyes: Independence-Interdependence | -0.06 | 0.12 | -0.53 | 0.59 |
| Condition Camera: Independence-Interdependence | -0.29 | 0.14 | -2.09 | 0.03* |

*Figure 4.13. Study 4. The influence of Independent-Interdependent Self-Construal on Destruction Rates in the Joy of Destruction Game.*



#### 4.4.2.4. Social Value Orientation as Moderator of Surveillance Cue Effects on Destruction Rates

In study 4 there were no significant interaction effects between SVO Scores and surveillance condition on destruction rates (Table 4.20).

Table 4.20. Study 4. Poisson GLM Examining Destruction rates in the JoD by Condition, SVO Score

| | *B* | *SE B* | *Z* | *P* |
|---|---|---|---|---|

Table 4.20. Study 4. Poisson GLM Examining Destruction rates in the JoD by Condition, SVO Score

| | | | | |
|---|---|---|---|---|
| Constant | 0.91 | 0.16 | 5.85 | <.001*** |
| Condition: Eyes | -0.27 | 0.21 | -1.28 | .20 |
| Condition: Camera | -0.18 | 0.23 | -0.78 | .44 |
| SVO Score | 0.001 | 0.003 | 0.31 | .58 |
| Condition Eyes: SVO Score | 0.001 | 0.003 | -0.82 | .76 |
| Condition Camera: SVO Score | -0.002 | 0.004 | -0.82 | .41 |

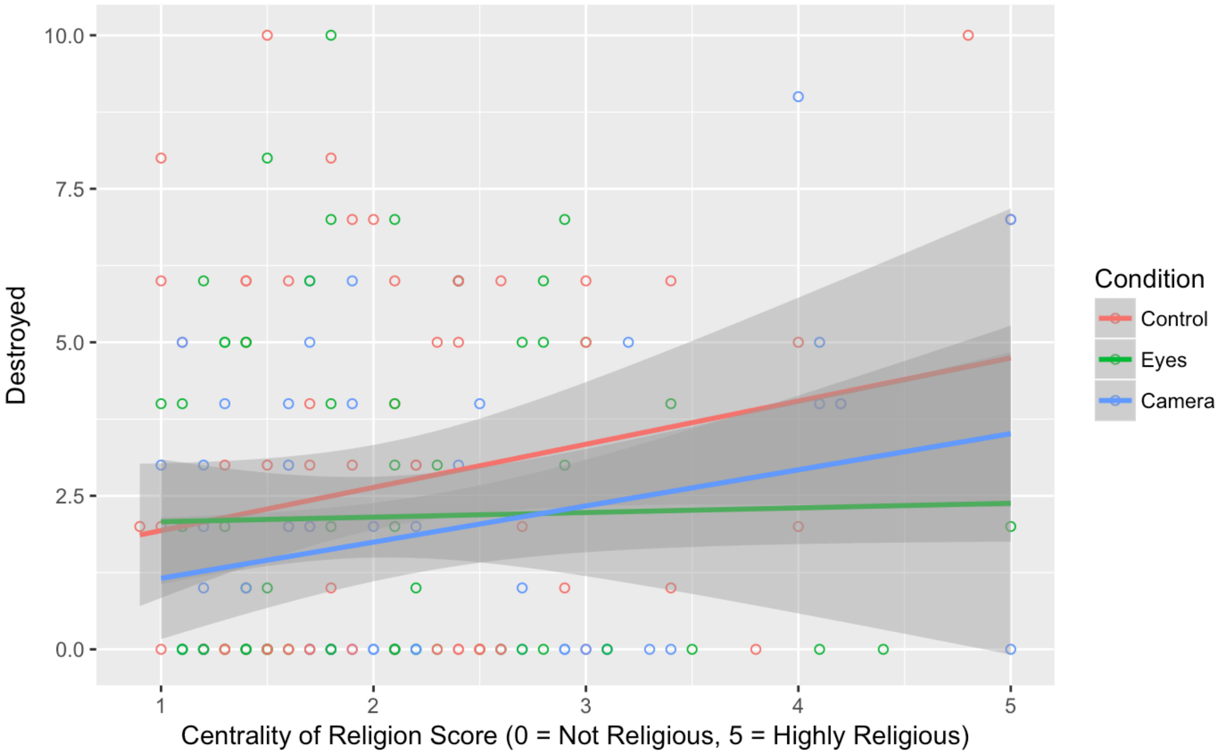
### 4.4.2.5. Religiosity as Moderator of Surveillance Cue Effects on Destruction Rates

In study 4, there was a significant effect of centrality of religion score on destruction ($b = 0.03$, $SE = 0.07$, $p < 01$), but no significant interaction by condition (see Table 4.21). Higher CRS Scores predicted increased destruction in the control and camera condition but not in the control as Figure 4.14 shows.

Table 4.21. Study 4. Poisson GLM Examining Destruction rates in the JoD by Condition, Religiosity

| | *B* | *SE B* | *Z* | *P* |
|---|---|---|---|---|
| Constant | 0.49 | 0.18 | 2.69 | <.01** |
| Condition: Eyes | 0.21 | 0..28 | 0.73 | .47 |
| Condition: Camera | -0.51 | 0.30 | -1.67 | .09 . |
| Centrality of Religion Score (CRS) | 0.23 | 0.07 | 3.09 | .002** |
| Condition Eyes: CRS | -0.20 | 0.12 | -1.58 | .11 |
| Condition Camera: CRS | 0.04 | 0.11 | 0.38 | .71 |

*Figure 4.14. Study 4. The influence of Centrality of Religion on Destruction Rates in the Joy of Destruction Game.*



### 4.4.3. Study 4 Discussion

The absence of any difference in destruction rates between Studies 2, 3 & 4 suggests that reducing perceive anonymity, or psychological distance, by removing the dice roll in the calculation of participant outcomes had no effect. This contradicts the findings of Abbink and Sadrieh (2009) and Abbink and Herrmann (2011) who found that such a change did reduce destruction rates. Similarly, we again failed to replicate Baillon et al. (2013) finding no influence of eye cues on destruction decisions. Again, as with studies 2 & 3, we failed to replicate Pfattheicher and Keller (2015), finding no influence of eye cues on public self-consciousness scores.

Looking for interaction effects, we found that as interdependence increased, destruction decreased in the camera condition, mirroring the effect found in Study 2, whereas in Study 3 there was no effect. This finding begins to suggest that the presence of a camera may predict lower destruction rates as individual's interdependence scores increase.

In study 4 as religion increased, destruction increased in the control and camera condition, but not in the eyes condition. Study 2 also found that increasing religiosity interacted with the presence of a camera to predict higher destruction but found that increasing religiosity predicted decreased destruction in the eyes condition. In the control condition the interaction effects were reversed between study 2 and study 4: increasing religiosity predicted lower destruction in the control in study 2, but higher destruction in study 4.

Since there were no significant differences in destruction rates, our primary dependent variable in our 3 JoD studies, we aggregated our results to explore the effects through the larger combined sample size of $N$ = 592.

## 4.5. Aggregate Analysis

### 4.5.1. Aggregate JOD Game Results

Table 4.22 compares studies 2, 3 & 4. We can see that there were no significant differences in either destruction rates, $H(2) = 1.35$, $p = .51$, nor in public self-consciousness, $H(2) = 1.49$, $p = .48$, between studies 2, 3 and 4. Comparing individual difference measures across the three experiments there was a significant reduction in independent-interdependent self-construal scores between experiments. Participants were less interdependent in Study 3 (*Mdn* = - 0.16) than in Study 2 (*Mdn* = 4.68), but as seen in Table 4.22, not significantly so. Table 4.22 also shows that in Study 4, participants were less interdependent still (*Mdn* = - 0.27) and that this difference was significant. Participants' SVO scores were significantly lower in Study 4 from in Studies 2 and 3, an increase accounted for by an increase in the number and proportion of individualists in Study 4 (29% of participants in Study 4 self-reported as individualists, a total of 56, whereas in Study 2 the figures were 4.4%, a total of 9 participants, and in Study 3 the proportion was 4.1% a total of 8 participants. Whether this was due to failed randomisation or was a real difference induced by condition cannot be known, but since our primary dependent variables – destruction and public self-consciousness did not differ between experiments we were justified in analysed our data in the aggregate across all three experiments.

Table 4.22. Summary Statistics Comparing Studies 2-4

| | | Study 2 | | Study 3 | | Study 4 | | Test Statistic | *P* | Test |
|---|---|---|---|---|---|---|---|---|---|---|
| *N* | | 206 | | 194 | | 192 | | | | |
| Male/Female/Unknown | | 137/64/5 | | 91/102/1 | | 104/87/1 | | | | |
| Destruction (0-1) | median [*IQR*] | 0.10 | [0.00, 0.50] | 0.20 | [0.00, 0.50] | 0.10 | [0.00, 0.43] | *H* = 1.35 | .51 | K-W |
| Expected Destruction | median [*IQR]* | 0.40 | [0.00, 0.50] | 0.40 | [0.00, 0.50] | 0.30 | [0.00, 0.50] | *H* = 0.72 | .70 | K-W |
| Public Self-Consciousness | median *[IQR*] | 11.00 | [6.00, 20.00] | 10.00 | [6.00, 15.00] | 10.50 | [7.00, 16.00] | *H* = 1.49 | .48 | K-W |
| Independence-Interdependence | mean (*sd*) | 4.68 | (0.72) | -0.16 | (0.88) | -0.27 | (0.93) | *F* = 8.91 | < .001** | ANOVA |
| SVO Score | median [*IQR*] | 44.00 | [42.05, 45.00] | 44.00 | [40.46, 45.00] | 28.24 | [18.15, 34.88] | *H* = 246.64 | < .001 | K-W |
| SVO Cat | Prosocial (%) | 196 | (95.1) | 186 | (95.9) | 132 | (68.8) | | | |
| CRS Score | median [*IQR*] | 1.80 | [1.50, 2.40] | 1.90 | [1.40, 2.80] | 1.80 | [1.40, 2.50] | *H* = 0.90 | .34 | K-W |

Values presented refer to the median, inter-quartile range, p-value and significance test associated with each variable.

Table 4.23 provides summary statistics comparing our aggregated results from Studies 2, 3 and 4. In the aggregate across experiments destruction rates were again significantly non-normal both by condition, Control $W = 0.83$, $p < .001$, Eyes $W = 0.84$, $p < .001$, Camera, $W = 0.81$, $p < .001$ and across all conditions combined, $W = 0.82$, $p < .001$. The assumption of homogeneity of variance in destruction rates by condition was also violated, $F(2, 589) = 3.78$, $p = .02$.

Aggregating across all three studies, median destruction rates (Table 4.23.) did not differ significantly by condition, $H(2) = 2.50$, $p = .29$.

Table 4.23. JoD Results Aggregated Across All 3 Studies, Stratified by Condition

| | | Control | | Eyes | | Camera | | Test Statistic | *P* | Test |
|---|---|---|---|---|---|---|---|---|---|---|
| *N* | | 216 | | 199 | | 177 | | | | |
| Male/Female/Unknown | | 116/96/2 | | 107/90/0 | | 109/67/0 | | | | |
| Destruction (0-1) | median [*IQR*] | 0.20 | [0.00, 0.50] | 0.10 | [0.00, 0.50] | 0.10 | [0.00, 0.40] | *H* = 2.50 | .29 | K-W |
| Expected Destruction | median [*IQR*] | 0.30 | [0.00, 0.50] | 0.30 | [0.00, 0.50] | 0.30 | [0.00, 0.50] | *H* = 0.01 | .99 | K-W |
| Public Self-Consciousness | median [*IQR*] | 10.50 | [7.00, 19.00] | 10.00 | [6.00, 15.00] | 11.00 | [6.00 19.00] | *H* = 3.74 | .15 | K-W |
| Independence-Interdependence | mean (*sd*) | -0.07 | (0.95) | -0.08 | (0.92) | -0.18 | (0.86) | *F* = 0.87 | .42 | |
| SVO Score | median [*IQR*] | 40.86 | [33.44, 45.00] | 43.14 | [31.56, 45.00] | 41.97 | [32.40, 44.00] | *H* = 1.00 | .61 | K-W |
| SVO Category | Prosocial (%) | 192 | (88.9) | 170 | (85.4) | 152 | (85.9) | | | |
| CRS Score | median [*IQR*] | 1.90 | [1.50, 2.60] | 1.80 | [1.40, 2.55] | 1.90 | [1.50, 2.50] | *H* = 0.91 | .64 | K-W |

Values presented refer to the median, inter-quartile range, p-value and significance test associated with each variable.

Logistic regression showed that in the aggregate, as in the individual studies, the model was no better than chance in predicting whether participants would destroy or not, Model $\chi^2(2)$ 1.09, $p$ = .58 and consequently there were no significant differences in the probability of destroying or not by condition (Eyes $z$ = -0.76, $p$ = .45; Camera $z$ = -0.28, $p$ = .79). The fit of our model was close to zero e.g. Hosmer-Lemeshow $R^2$= .001 suggesting the presence or absence of surveillance cues is a very poor predictor of whether participants choose to destroy. The aggregated odds ratios suggest that in fact the presence of eye cues predicted an *increase* in destruction of c.16%, though the confidence interval is very wide and crosses zero, *OR* =1.16, *95% CI* = 0.79, 1.72. The odds ratio in the camera condition suggested that camera presence predicted a decrease in destruction of c.6% but again the confidence interval was very wide, *OR* =.94, *95% CI* = 0.63, 1.41. The results of the logistic regression are in Table 4.24.

Table 4.24. Logistic Regression of data aggregated across Studies 2-4 Condition vs Destroyed or Not

| | | 95% CI for odds ratio | | |
|---|---|---|---|---|
| | *B(SE)* | Lower | Odds Ratio | Upper |
| Included | | | | |
| Constant | 0.20<br>(0.14) | | | |
| Eyes | 0.15<br>(0.20) | 0.79 | 1.16 | 1.72 |
| Camera | -0.06<br>(0.20) | 0.63 | 0.94 | 1.41 |

Note. $R^2$= .001 (Hosmer-Lemeshow), .002 (Cox-Snell), .002 (Nagelkerke), Model $\chi^2(2)$ 1.09, $p$ = .58

As we can see from the four graphs at figure 4.15, most participants did not act antisocially/maliciously, choosing not to destroy regardless of condition. The next most popular strategy involved choosing to destroy 50% of the time. Our data demonstrates that most participants made at least one malicious decision in circumstances where there was no benefit to them in doing so, but that there was no effect of our surveillance variables on such behaviour. Figure 4.16. shows that participants destroyed in proportion to the extent to which they expected others to destroy their allocation.

*Figure 4.15. Distribution (%) of Decisions to Destroy by Participant in the Joy of Destruction Task*

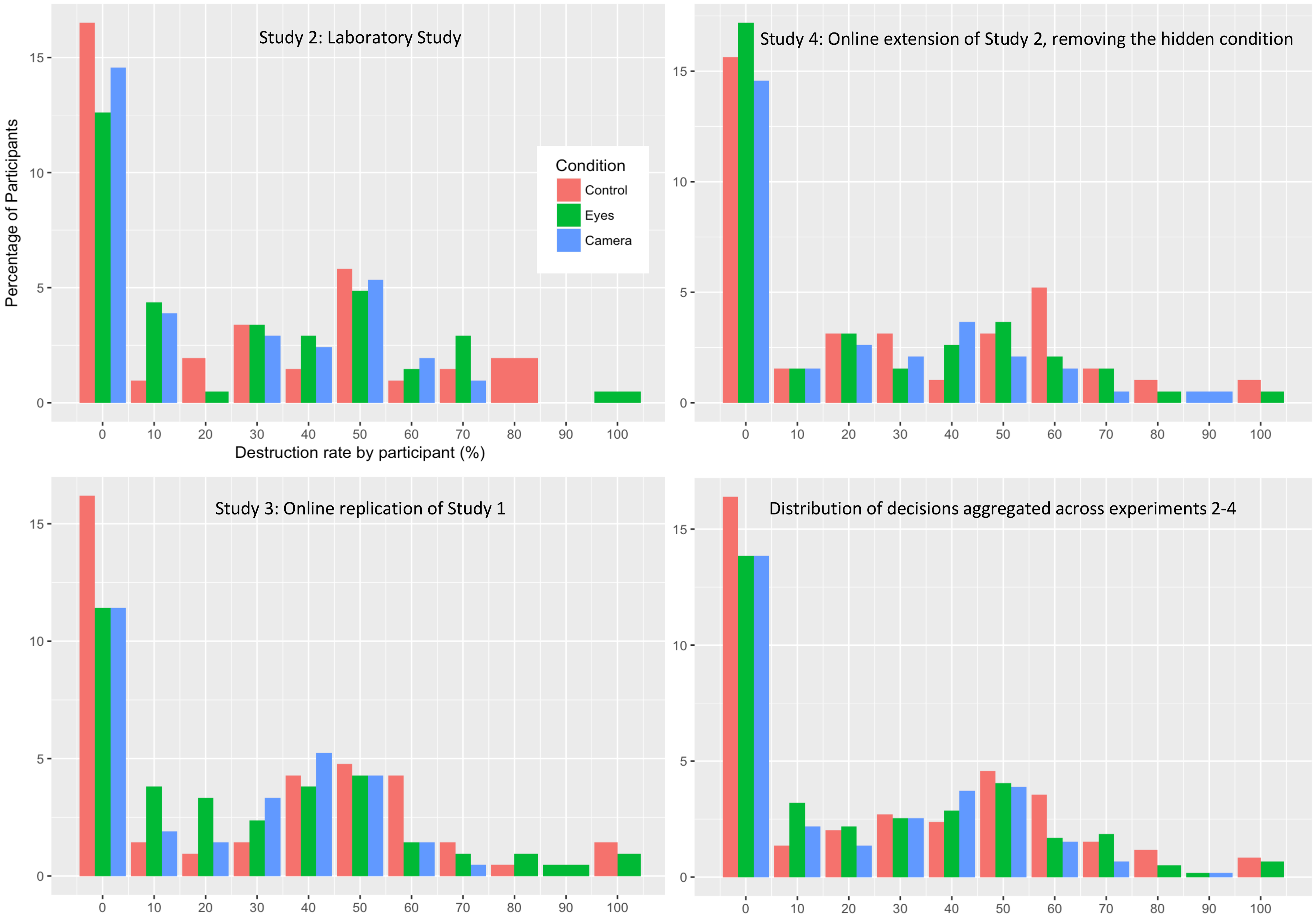
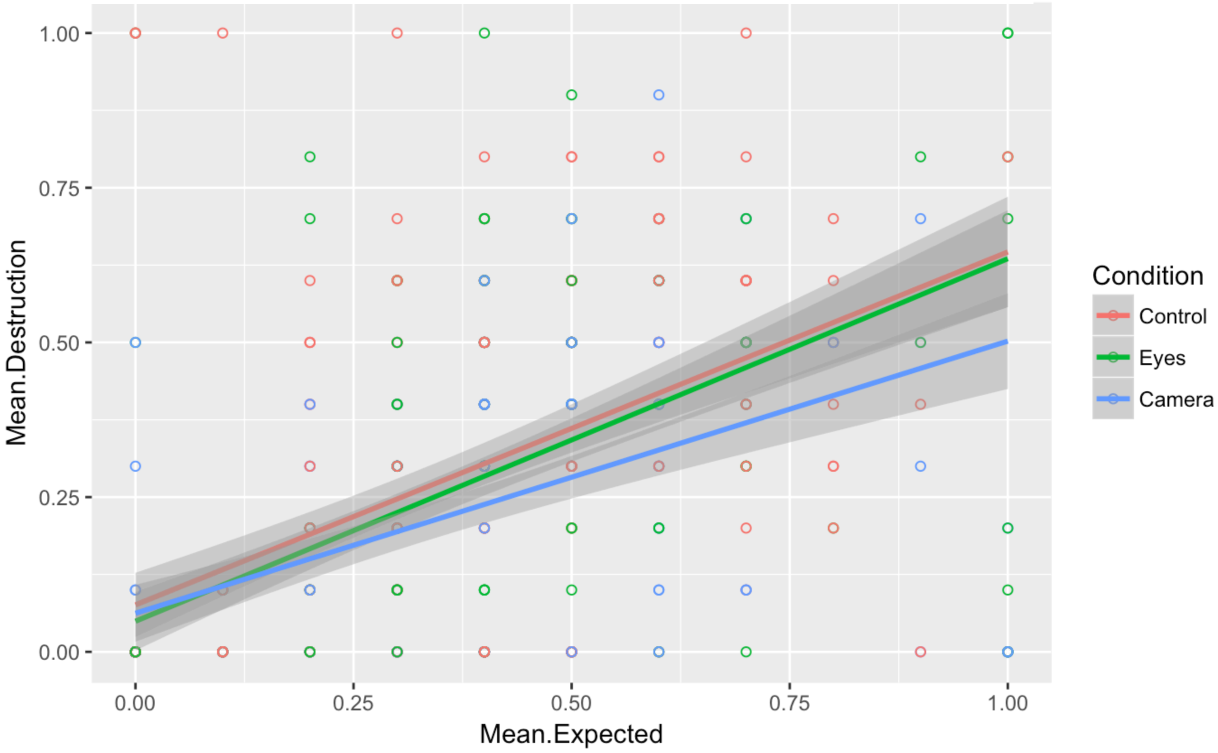


*Figure 4.16. Aggregate Analysis: Mean Expected Destruction by Mean Destruction*



### 4.5.2. Public Self Consciousness and Surveillance Cues

Public self-consciousness in the Eyes (*N*=199, *Mdn*=10.00, *IQR* = 6.00, 15.00) and Camera (*N*=177, *Mdn* = 11.00, *IQR* = 6.00, 19.00) conditions, did not significantly differ from self-reported public-self-consciousness in the Control (*N* = 216, *Mdn* =10.50, *IQR* = 7.00, 19.00), *H(2) = 3.47, p* = .15 (see Figure 20)

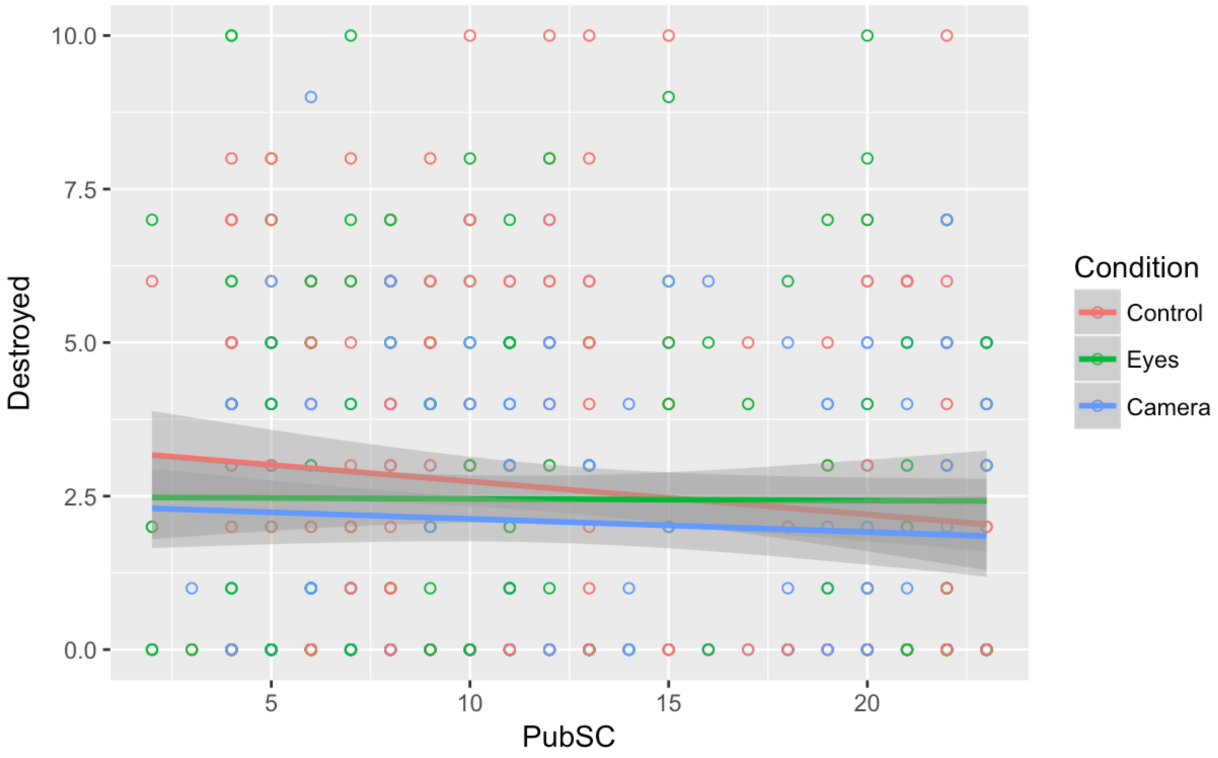
### 4.5.3. Public Self Consciousness as Moderator of Surveillance Cue Effects on Destruction Rates

Our Poisson GLM (Table 4.25.) explored whether those that were high in public self-consciousness were more influenced by surveillance cues. We predicted that in the eyes and camera conditions the higher a participants' public self-consciousness score the less they would destroy, whereas in the control condition there should be no difference between participants, or perhaps high public self-consciousness individuals would destroy more. The main finding from our aggregate results is clear in the scatterplot at Figure 4.17. There was little meaningful difference between destruction rates by condition and little to no meaningful interaction effect with public self-consciousness. Participants destroyed consistently around 2-3 times of the maximum 10 possible. The overlapping confidence intervals and the contradictory direction of the significant effects found in Study 2 when compared with Study 3, the absence of significant effects in Study 4 and the increasing convergence of the betas by condition in our aggregate analysis suggests we should be cautious in interpreting these significant results as meaningful.

Table 4.25. Aggregate Analysis. Poisson GLM Examining Destruction rates in the JoD by Condition, Public Self-Consciousness

| | *B* | *SE B* | *Z* | *P* |
|---|---|---|---|---|
| Constant | 1.21 | 0.09 | 13.85 | <.001*** |
| Condition: Eyes | -0.30 | 0.13 | -2.37 | .02 * |
| Condition: Camera | -0.36 | 0.14 | -2.51 | .01 * |
| Public Self-Consciousness | -0.02 | 0.01 | -3.13 | .002 ** |
| Condition Eyes: Public Self-Consciousness | 0.02 | 0.01 | 2.01 | .04* |
| Condition Camera: Public Self- Consciousness | 0.01 | 0.01 | 1.01 | .31 |

*Figure 4.17. Aggregate Analysis. The influence of Public Self-Consciousness (PubSC) on Destruction Rates in the Joy of Destruction Game.*



### 4.5.4. Interdependence Scores and Surveillance Cues as Moderator of Surveillance Cue Effects on Destruction Rates

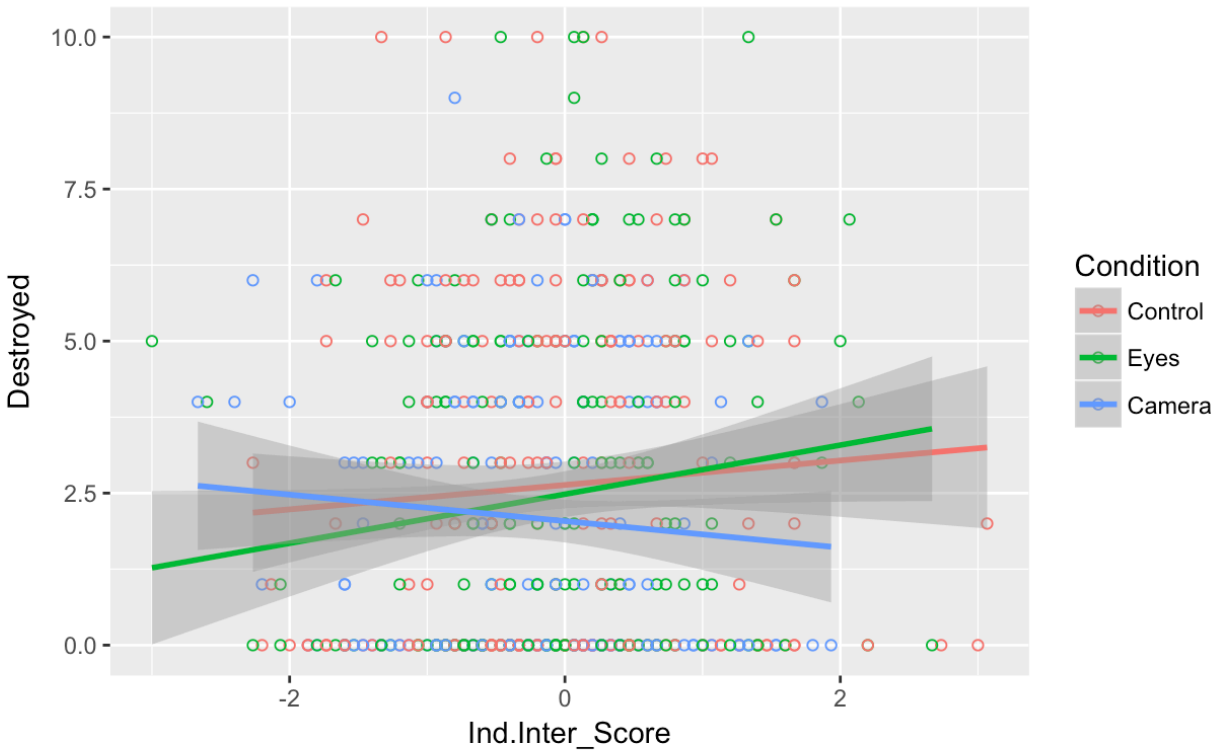
Table 4.26 shows there was a significant interaction between condition, destruction rates and participants' score on the continuous independent-interdependent self-construal scale.

Table 4.26. Aggregate Analysis. Poisson GLM Examining Destruction rates in the JoD by Condition, Independence-Interdependent Self-Construal.

| | *B* | *SE B* | *Z* | *P* |
|---|---|---|---|---|
| Constant | 0.97 | 0.04 | 22.98 | <.001<br>*** |
| Condition: Eyes | -0.07 | 0.06 | -1.10 | .27 |
| Condition: Camera | -0.26 | 0.07 | -3.74 | <.001*** |
| Independence-Interdependence | 0.08 | 0.04 | 1.72 | .09 . |
| Condition Eyes: Independence-Interdependence | 0.09 | 0.07 | 1.33 | .18 |
| Condition Camera: Independence-Interdependence | -0.18 | 0.07 | -2.40 | .02 * |

Figure 4.18 shows that in the Control and Eyes conditions the more interdependent participants destroyed more than those who were more independent, as we would predict. We had also predicted that there would be an interaction between eye cues, interdependence scores and destruction rates but there was not.

*Figure 4.18. Aggregate Analysis. The influence of Independent-Interdependent Self-Construal on Destruction Rates in the Joy of Destruction Game.*



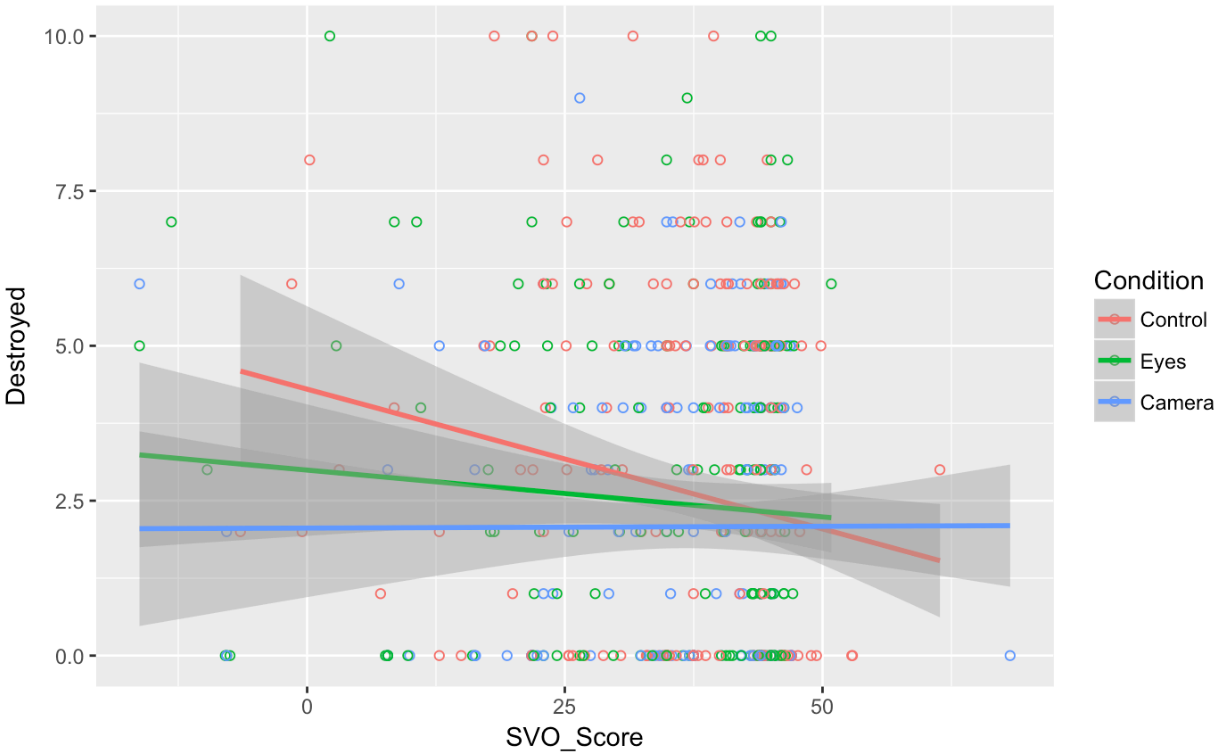
### 4.5.5. Social Value Orientation and Surveillance Cues as Moderator of Surveillance Cue Effects on Destruction Rates

In the aggregate analysis there was a main effect of SVO Score on destruction rates with high prosocials destroying less. The significant interaction effects between both surveillance cues and SVO scores was driven by the stronger association between increasing prosociality and decreasing destruction in the control condition. (Table 4.27, Figure 4.19).

Table 4.27. Aggregate Analysis. Poisson GLM Examining Destruction rates in the JoD by Condition, SVO Score

| | *B* | *SE B* | *Z* | *P* |
|---|---|---|---|---|
| Constant | 1.51 | 0.13 | 12.00 | <.001*** |
| Condition: Eyes | -0.41 | 0.17 | -2.42 | .02* |
| Condition: Camera | -0.79 | 0.21 | -3.75 | <.001*** |
| SVO Score | -<br>0.015 | 0.003 | -4.49 | <.001*** |
| Condition Eyes: SVO Score | 0.01 | 0.004 | 2.06 | .04* |
| Condition Camera: SVO Score | 0.02 | 0.01 | 2.78 | .005** |

*Figure 4.19. Aggregate Analysis. The influence of Social Value Orientation on Destruction Rates in the Joy of Destruction Game.*



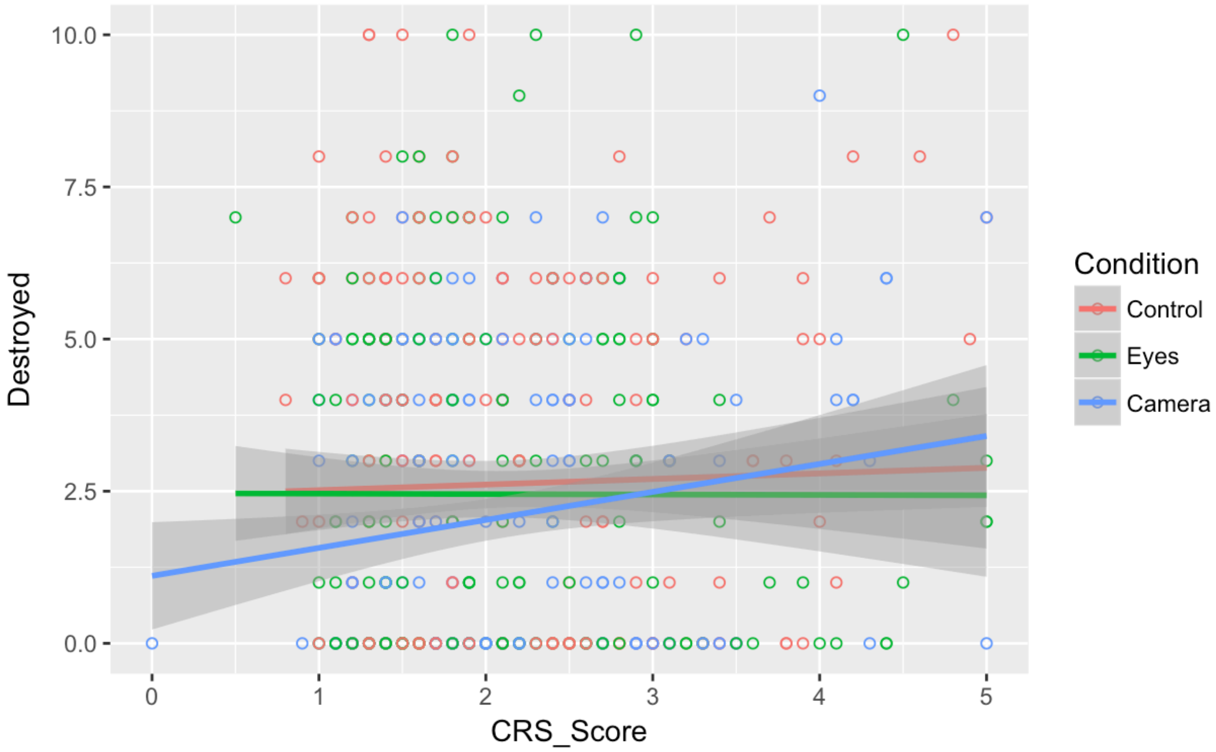
### 4.5.6. Religiosity and Surveillance Cues as Moderator of Surveillance Cue Effects on Destruction Rates

There was no significant interaction between the eye cues condition, centrality of religion scores and destruction frequency (Table 4.28). However, there was a significant interaction between the camera condition, participants' centrality of religion scores and destruction. Figure 4.20 shows that in the camera condition the more religious a participant, the more they destroyed, low religious participants destroyed significantly less frequently than similarly low religious participants in the control, while the highly religious destroyed more frequently in the presence of a camera than similarly highly religious participants in a no-camera control. The correlation between greater centrality of religion in a participants' life and higher destruction rates in the camera condition was consistent across all 3 studies. There was no consistent interaction effect between eye cues and condition on destruction rates.

Table 4.28. Aggregate Analysis. Poisson GLM Examining Destruction rates in the JoD by Condition, Religiosity

| | *B* | *SE B* | *Z* | *P* |
|---|---|---|---|---|
| Constant | 0.89 | 0.11 | 8.04 | <.001*** |
| Condition: Eyes | 0.01 | 0.16 | 0.08 | .93 |
| Condition: Camera | -0.60 | 0.17 | -3.45 | <.001*** |
| Centrality of Religion Score (CRS) | 0.03 | 0.05 | 0.73 | .47 |
| Condition Eyes: CRS | -0.04 | 0.07 | -0.53 | .60 |
| Condition Camera: CRS | 0.17 | 0.07 | 2.32 | .02* |

*Figure 4.20. Aggregate Analysis. The influence of Centrality of Religion on Destruction Rates in the Joy of Destruction Game.*



## 4.6. Discussion

Across three experiments we found that participants were no more self-conscious in the presence of eye cues or a camera nor were they so when the data was analysed in the aggregate. This is in contrast to the findings of Pfattheicher and Keller (2015), whose insight – that measuring public self-consciousness might be a good proxy for measuring how much participants felt watch – drove the inclusion of public self-consciousness scores in our study. Pfattheicher and Keller's finding provided the best self-report measure that has yet been found to support the idea that watching eyes effects could be specifically attributed to making people feel more watched. Experiments that find the presence of eye cues changed behaviour universally explain this effect by suggesting that the watching eyes cued an automatic mechanism that make people feel watched and activated implicit reputation management mechanisms (e.g. Haley & Fessler, 2005; Mifune, Hashimoto, & Yamagishi, 2010; Vaish, Kelsey, Tripathi, & Grossmann, 2017). However, our failure to replicate Pfattheicher and Keller's (2015) findings that eye cues reduce anti-social behaviour heightens questions within the field as to whether there really is a watching eyes effect. For those who contend that there is, our failure to replicate leaves the field in need of greater evidence both for how they do so. Our interaction effects between surveillance cue and self-consciousness scores were in the opposite direction to that hypothesised, were inconsistent, and absent both in Study 3 and the more highly powered aggregate analysis.

In none of our three experiments, nor in our analysis of all three in the aggregate did we find any convincing evidence for the effect of watching eyes on malicious behaviour. Thus, in our three experiments we failed to replicate Baillon et al.'s (2013) findings that eye cues reduced anti-social behaviour in the Joy of Destruction game. Each of our individual experiments employed marginally more statistical power than Baillon et al., each was more carefully designed to elicit a watching eyes effect by using the posters incorporating a picture of watching eyes from the papers by Nettle, Bateson and colleague at Newcastle University (rather than the image of the statue of Erasmus selected by Baillon and colleagues). Both individually, and in aggregation, our studies failed to replicate any watching eyes effect on malicious behaviour. Of course, it is possible that such an effect does exist as we explore in looking at the limitations of our research and suggestions for future studies. However, the most parsimonious explanation is that there is no effect of watching eyes on malicious behaviour in the JoD, with Baillon et al the outlier as a false positive.

We also failed to replicate Abbink & Sadrieh (2009) and Abbink & Hermann (2011). Abbink & Sadrieh and Abbink & Hermann showed that even when people were anonymous they rarely acted maliciously (destruction rates average 8.5% and 11% respectively), suggesting that this was down to their concern for the psychological costs inflicted on the victim who would know that another human had inflicted harm on them. However, once participants could hide behind the possibility that chance had inflicted harm on another such that their victim would be uncertain as to whether they were unlucky or whether they had been the subject of a malicious act, participants were twice as willing to inflict malicious harm. Abbink & Hermann suggest this is because there is a pleasure to be derived from being nasty, or a joy to be taken in destruction. However, our results showed that destruction rates remained consistent regardless of whether decisions were ‘hidden’ or not. Since destruction rates did not differ significantly by condition we take destruction rates in the control condition as our illustrative baseline. They showed little variation between studies. In Study 2 and 3, when chance ‘hid’ decisions, the mean destruction rate in the control was respectively 24% and 28%. In Study 4, with the role of chance removed and decisions no longer ‘hidden’, the destruction rate was 26%.

We found no interaction between the eyes condition and the independence-interdependence scale, failing to replicate Kitayama et al.’s (2004) findings that independent participants destroyed more in the presence of social cues while more interdependent participants destroyed less. However, we did find, in Studies 2, 4 and in the aggregate analysis, an interaction effect between the presence of a camera and participants' independence-interdependence score that aligned with Kitayama's prediction: the more independent destroyed more when a camera was present than in a control, the more interdependent destroyed less with a camera present than in its absence. We should note that Study 3 found the opposite, although the interaction was not significant. What can we conclude? We might tentatively advance the hypothesis that surveillance cameras induce less malicious behaviour amongst the more interdependent, and greater malicious behaviour amongst the more independent. With greater empirical support such a finding would have important policy implications, suggesting that in interdependent societies such as those in China and Japan camera surveillance would have broadly beneficial effects, reducing crime or fraud for example – much as Bentham predicted. In more independent societies it might be actively counter-productive. But we should be cautious of placing too much weight on findings that while statistically in significant in 2 of 3 studies as well as in the aggregate analysis of all 3 studies taken together, also showed overlapping confidence intervals in all studies and were contradicted in one of the three studies.

We found no consistent interaction effect of participants' social value orientation and their centrality of religion score on destruction by condition. We thereby failed to show, as Luo et al. (2016) had previously, that SVO score interacted with watching eyes to increase cooperative behaviour.

Nor did we find any evidence to support the supernatural surveillance hypothesis. Where the supernatural surveillance hypothesis would predict no difference in destruction rates between the control and surveillance cue conditions for the highly religious participants – since they always feel watched by God – we found conflicting effects in each study. In the aggregate we found no difference in destruction rates for low and highly religious participants in the control and eyes conditions, while in the camera condition the low religious destroyed less in the absence of a camera than low religious participants did in the control, and the highly religious destroyed more in the presence of a camera than in the control. Thus, the effect of centrality of religion was highly inconsistent and cannot be said to have meaningfully influenced destruction rates by condition.

Our broad replication of previous Joy of Destruction findings – that people will act maliciously even when there is no incentive to do so - is notable. Some 44% (N=261) of participants followed the income-maximising and non-malicious and rational path of never choosing to pay to reduce a partner's income. Over half our participants (56%, N = 331) chose to act maliciously at least once, and on average 24% (N=592) of decisions made were malicious. These findings seem at face value to be either profoundly irrational, or perhaps to indicate that there is something sinister in many of us - a pleasure to be found in being nasty, to paraphrase the title of Abbink & Sadrieh's (2009) paper that introduced the JoD task.

Brañas-Garza, Espín, Exadaktylos, and Herrmann (2014) find evidence beyond the dictator game to suggest there exists a consistent number of people that will behave maliciously, finding 17% of participants in an ultimatum game that made unfair offers and rejected fair offers. This evidence in turn aligns with that documented by Herrmann, Thöni, and Gächter (2008) in participant pools across 16 different nations, with anti-social punishment consistently present. But this may be an over-simplification. Leibbrandt and Lopez-Perez (2011) showed that inequality aversion can be so extreme that third parties will pay to punish even those who always act fairly, if the net result of not doing so would be that that fair individual ends up with larger share of the collective wealth than those they have interacted with.

This apparent envy or resentment leading people to punish even those who have done no wrong, might explain actions within the JoD. It is possible that participants feared that others might be maliciously reducing their income, and wished to ‘even the score’. Indeed Cox & Li (2012) have shown that in economic games obfuscating whether someone has been betrayed or not increases their tendency to betray, compared to when decisions are transparent. Thus, inequality aversion might in part explain our findings. If so, it is notable that expected destruction was consistently higher than actual destruction – suggesting participants were willing to tolerate some inequality but tried to insure themselves against being too naïve.

Our survey ended with an exit question which asked participants why they made the decisions they had in the first task (the JoD game). Responses suggest that at least for some, inequality aversion was the motivation and fear of exploitation was a motivating factor, while for others, actions were based on the golden rule of ‘not doing unto others as you would not have done unto you’.

It is possible that the task used here lacks ecological validity. Although paying a cost to punish is irrational when viewed from the deep logic of evolution – reducing one’s fitness for no gain - proximate mechanisms might explain it. As previously mentioned Krasnow & Delton (2016) argue that deep evolutionary logic renders our love of dense, fatty, sugary, high calorie foods rational and adaptive even while the expression of this love in the modern context renders us obese and the behaviour damaging. Similarly, they suggest that the male desire for females taking the contraceptive pill is not adaptive but is triggered by proximate mechanisms – the desire to produce offspring - that were adaptively rational in our early evolutionary context. Both are examples of proximate mechanisms misfiring while the deep evolutionary logic remains sound. In the JoD, the absence of feedback on others decisions lacks ecological validity, and thus destruction may reflect the misfiring of a proximate mechanism that has evolved to expect others to defect some of the time and thus causes us to punish to reduce defection. In this understanding, destruction becomes rational from the perspective of the proximate logic of the evolutionary mechanism.

Finally, if we consider human decision-making at a collective, or a systems level, we should consider the import of errors, noise, non-conformity or dissent as an explanation for the decisions to destroy. Shirado and Christakis (2017) have shown that the introduction of 'noisy agents', bots that produce random decisions, can improve human coordination and help resolve collective action problems. In their study, bots that made random decisions alongside optimal ones at a rate of 10% helped increase problem solving rates and speed. 85% of those networks with the 'noise' solved the problems within 5 minutes, whereas only 67% of those without any noise were resolved within the same time frame. Such models may have utility in resolving local-global coordination problems in real-world settings (Gächter, 2017). It might be that groups require the noise of decisions that go against the consensus. Thus, whether individuals were motivated to destroy by obdurate non-conformity (an inverse form of experimenter demand effects), the desire to dissent or just to create some 'noise', at a systems level these decisions might have met a deeper evolutionary imperative that requires random individual errors to improve cooperation and enhance collective performance. Further evidence for this idea comes from Efferson et al (2008). Efferson found in his economic game that 30% of participants left money on the table rather than conform, suggesting that perhaps these are the "mavericks" that social innovation needs – people who refuse to conform. Maybe the JoD is about divergence, not just influence or maliciousness?

## 4.7. Limitations

The Joy of Destruction game conflates prosocial and selfish actions. To illustrate: if I am ruthlessly self-interested and rational I don't destroy in order to maximise my income.

Similarly, if I am deeply altruistic I don't destroy in order to maximise the other's income. And again, if I am aim to balance my self-interest with that of others, that is if I am committed to fairness, I don't destroy in order to maximise mutual incomes. In a sense this is unavoidable as we want to identify, isolate and test surveillance cues' effects on malicious behaviour as the least contestable form of anti-social behaviour. It is reasonable to suggest that *if* eye cues make us feel watched they might influence the selfish the most – those who would maximise their own advantage whenever they feel they can, but manage the impression they make on others carefully to ensure they do not lose the benefits of a good reputation. Perhaps next most affected might be altruists and the prosocial. In Chapter 3 we discussed at length the implications for the watching eyes effect of research suggesting that generosity is not always good for one's reputation, and extreme generosity can be bad for it (Klein & Epley, 2014; Klein, Grossmann, Uskul, Kraus, & Epley, 2015). If this is so, altruists and the prosocial might be more generous when anonymous, but less so as anonymity is decreased – therefore more influenced by eye cues than the actively malicious. In our study, altruists and the prosocial, like the selfish, would not destroy at all in order to maximise others' incomes. The design of the Joy of Destruction task isolates the malicious at the expense of not being able to test for differences in the behaviour of the selfish, prosocial and altruistic.

The significant correlation between expected destruction and the decision to destroy also raises questions as to whether we can describe these decisions as malicious. If participants destroy because they think others will do the same perhaps they can be said to be seeking income equality and fairness in outcomes for both parties through mirroring their partners' expected decisions. Similarly, we might consider this in light of Fenigstein & Abrams (1993) – who found that the presence of a camera increased people's sense that others shared their outlook on the world 'a false consensus effect' – it might be that participants listened to this worry more in the presence of eye cues and cameras, off-setting the cueing of reputational worry that they might be perceived as acting maliciously with the fear they might being seen as naïve and exploitable.

Relatedly, it is possible that people were motivated to destroy not by a pleasure to be found in being nasty but rather by the desire for influence. A recent laboratory-based experiment demonstrated that people were willing to pay for influence in simulated social settings even when that influence brought them no personal gain (Hertz, 2018). Sadrieh and Schröder (2012), tested the hypothesis that altruists and the malicious might not be separate groups at the opposite end of an anti-to-pro social spectrum, but rather that they might just be one group with a desire to influence others. They ran public goods games in which it was possible to destroy the income of others, to altruistically give to others or to maximise payoff to self. They found that one third of participants were those with 'a desire to influence others' and would give or destroy to do so. A further third were payoff maximisers, and a final third were solely prosocial. This finding suggests the JoD might be confounding the desire to influence with malicious behaviour since in our task the two are the same.

An extension of this idea, that people in our task might have been more interested in influence than in hurting others, can be found if we consider effectance motivation. Described in a theory first put forward by R. W. White (1959) effectance motivation describes the human drive to manipulate and experiment with elements in the world in a 'cycle of transaction' involving stimulation for the sake of novelty, to deliver a feeling of efficacy, or discover the limits of our efficacy – thus people might have been choosing to destroy out of curiosity, just to see what happens rather than sitting 'doing nothing' throughout the task (See also: Harter, 1978; or for the related self-determination theory: Ryan and Deci, 2000).

A neat demonstration of effectance motivation can be found in a study by T. D. Wilson et al. (2014). They found that people's desire to influence extends beyond influencing others, to the desire to influence the environment even when that comes at a cost. First they asked participants to rate the pleasure or pain of a range of stimuli, such as viewing attractive photographs or experiencing electric shocks, before stating how much they would pay to experience or avoid the same stimuli again. Many said they would pay to avoid the electric shocks again. The participants were then put in isolation to see how they coped with 'being alone with their thoughts', with the electric shock equipment present. They were told they could use the equipment but there was no need to do so. Yet still, of the participants who said they would pay to avoid being shocked again, 67% of men and 25% of women did give themselves an electric shock. Many people preferred to be doing something rather than nothing, even if that something was negative. We might interpret this as a need for influence over others, our environment or ourselves, particularly when there is nothing else to do. White's effectance motivation and Wilson's empirical demonstration of it suggest that 'maliciousness' in our experiment might be confounded by the desire for influence of any kind over inactivity.

A further challenge lies in the difficulty of preventing people feeling watched in the laboratory. Participants know that their data is being collected. Therefore, regardless of written assurances of data anonymity, the enclosed nature of the testing booths and the fact that they are alone in the room, the whole experiment is framed in a way to heighten self-consciousness. All of which allude to the possibility that under such circumstances it becomes impossible to have a true control condition, where people genuinely don't feel watched. In a recent experiment designed to test this idea Noah, Schul, and Mayo (2018) interpreted differences in participants perceptual fluency – the length of time they estimated it would take them to prepare a recipe - between participants who checked their name off a list, signed a consent form and then were handed a questionnaire (standard experimental design), and those handed a questionnaire in an envelope without checking their name of an attendance list and without signing a consent form (conveying greater anonymity) - as suggesting that the standard baseline design of laboratory experiments might be enough to cue people to feel watched.

## 4.8. Further Research

Given that it may be that those most likely to be affected by eye cues, the altruistic and the selfish, were both unlikely to destroy within the design of the Joy of Destruction, it would be desirable to test the effect of eye cues and cameras on decisions when they can vary between generous, self-interested and malicious. Tasks that allow both altruism and maliciousness (e.g. Sadrieh & Schröder, 2012) to explore whether it is those seeking influence that are affected by surveillance cues rather than the generous or the malicious might help to better understand the boundary conditions of any surveillance cue effect, or to add to the evidence suggesting just feeling watched does not influence behaviour.

Future researchers might helpfully measure the dark triad personality traits of Machiavellianism, narcissism and psychopathy (Jones & Paulhus, 2013; Paulhus & Williams, 2002), which are widely used measures of malicious behaviour, and despite limitations are currently the best psychological measure we have for such behaviour (for a review see King et al., 2018). It would be useful to see whether the more malicious destroy more frequently than those low in maliciousness as a manipulation check to confirm that the JOD is testing maliciousness, as assumed in our studies and in prior uses of the task by other authors (e.g. Abbink and Sadrieh, 2009; Abbink & Hermann, 2009; Baillon, Selim and Van Dolder, 2013). It would also be valuable to understand if the more malicious are more or less affected by surveillance cues.

# Chapter 5

# 5

# Study 5 Watching Eyes on Crime in Hereford: A Collaborative Field Experiment

> *"For the eyes of the LORD run to and fro throughout the whole earth…"*
>
> 2 Chronicles 16:9

In a field study we examined the effect of signs bearing watching eyes (and other priming images) on crime over 7 months at 9 locations across the town of Hereford, West Mercia, UK. Additional priming cues included pictures of surveillance cameras and a police logo, as well as all-capitals text slogans such as 'CYCLE THIEVES WE ARE WATCHING YOU'. With the priming cues layered in this way it is not possible to isolate 'what worked' in our intervention. However, we can say that the presence of the signs correlated with statistically significant decreases in crime at 3 of 9 locations, and that crime or antisocial behaviour reduced at every site where signs were erected. Additionally, at 6 of 9 sites anecdotal reports from local 'guardians' suggested there had been much larger reductions in crime than we were able to record in police calls or independent measures. Most strikingly, at the 3 sites where data was most accurately collected, a major High Street clothing store, a hospital cycle shed, and a fly-tipping site we find a 27%, 100% and 100% reduction in crime following the installation of the signs.

# 5. Watching Eyes on Crime in Hereford

## 5.1. Introduction

This chapter reports the results of the first of two field studies exploring the effectiveness of using signs bearing images of eyes and other cues in reducing crime and anti-social behaviour. These results are presented in the context of the watching eyes literature, but it is important to note that this study, as with others such as the landmark Newcastle University study by Nettle et al. (2012), might better be described as an examination of layered priming effects. We applied three principle primes: a simple signage intervention, the presence of eye cues, and an authority prime via the inclusion of a police logo. We discuss the opportunities and challenges involved in running collaborative field experiments and offer some lessons for others engaging in such partnerships.

Chapter 3's systematic review and meta-analysis suggested that eye cues might reduce anti-social behaviour by as much as 35% . However, in three well-controlled studies with large sample sizes (Chapter 4) no effect of eye cues on anti-social behaviour were found. Since over half (8 of 15) of the experiments included in the meta-analysis were field experiments, it is possible that eye cue effects on anti-social behaviour are more consistent in the field. There is some evidence that make this a plausible explanation. Burnham (2003) posited that in laboratory experiments it is very difficult to create a situation in which a participant does not perceive their actions to be 'watched' in some way. Their anonymity might be assured in pre-experiment briefings but participants know their behaviour is being measured and recorded – that's the point of psychology experiments. Implicitly priming participants to feel watched is unintentionally and unavoidably intrinsic to all laboratory experiments. Additionally, some participants, perhaps most, will be aware that psychology experiments often involve deception that isn't revealed at the start. Consequently, assurances of privacy may not be enough to avoid implicit or explicit concerns to the contrary. Some researchers have eschewed the use of computers for much the same reason - since participants can never be sure exactly what the computer is recording in contrast to pen and paper studies (Bolton & Zwick, 1995). Consequently, we developed a collaborative partnership with West Mercia Police to instigate a crime reduction initiative testing the watching eyes effect on anti-social behaviour in a naturalistic environment.

It is not possible to isolate the effect of watching eyes from other cues in our design. Research in the 1970s showed that just the presence of a simple sign (text only) can be enough to reduce criminal behaviours (McNees, Egli, Marshall, Schnelle, & Risley, 1976), so it might just be the text in, or presence of, the sign that led to a reduction in anti-social behaviours. Our signs also incorporated prominent police and business logos. Cialdini's influential work on persuasion suggests that authority appeals, the invocation of authority, even if merely carried by a symbol, can help to change behaviours (Cialdini, 2007). Related if limited evidence supports his assertion in studies looking at the use of titles and authority appeals in email scams and on the effect of authority appeals conveyed via smart dress and uniforms (Archer, 2017; Bushman, 1984) and of course there is the foundational body of work on compliance to authority developed by Milgram in the 1960s and now well established in social psychology (T. Blass, 1991; Thomas Blass, 2012; Burger, 2009; Milgram, 1963, 1974, 1977; Reicher & Haslam, 2011; Stanley, 1974). Consequently, it might be that police and business logos caused people to be more aware of authority and thus improve their behaviours. Finally, the presence of a CCTV camera logo might have caused reductions in anti-social behaviour, since an observer might reasonably assume that they are being filmed by CCTV. That the presence of CCTV can, in certain limited and discrete circumstances, cause reductions in crime has been demonstrated in meta-analysis by Welsh and Farrington (2009) and others as described in Chapter 1. We must be careful when ascribing causality to any single element of our signs, since all aspects might have some influence based on prior research.

The competing imperatives of the researchers, that is our need for carefully balanced and crafted experimental design to isolate the effects of ‘watching eyes’ from potentially confounding factors, and those of the police service, that is the need to maximise the potential effectiveness of the intervention to minimise criminal and anti-social behaviours, required compromise. Consequently, ours was a layered priming effects study.

## 5.2. Method

### 5.2.1. Design

The dominant intervention was the picture of watching eyes, but also present was a stylized image of a CCTV camera, instructive text in red & white, and the various police, commercial and council logos.

### 5.2.2. Study Sites

The sites selected were chosen by Hereford Police. Where the signs were to be erected in commercial premises consultation with the store manager was necessary. The necessity of such consultation prevented the experiment being as controlled as we would have wished, since as least some of those reporting criminal or anti-social behaviour at McDonalds and the country clothing store would know of the intervention we had put in place. The same limitation applied to Hereford Cathedral where the small number of staff all knew the signs had been put up, and at Aylestone Park, where the residents association had to be informed. Similarly the gamekeeper at Huntsham Bridge and the caretaker at the school at the fly-tipping site had to give their permission for the signs to be put up and thus their reports were not independent of the intervention. In the case of the fashion store blinding was maintained as the analysis was run on stock loss data and not on calls to the police. The sites are listed at table 5.1. along with a brief description of the location, the type of behaviour we sought to affect, and the outcome measure used.

Table 5.1.

Intervention Site Listings

| Serial | Location | Description | Behaviour Targeted | Outcome measure | Additional Signs Present? |
|---|---|---|---|---|---|
| (1) | McDonalds | Fast food chain | Anti-social behaviour | Customer complaints & calls to the police | Y |
| (2) | Hereford Country Hospital | A large bike shed at Hereford Hospital | Cycle theft | Recorded cycle thefts | N |
| (3) | Country Clothing & Guns Store | Independent retailer | Stock theft | Stock loss figures at audit | N |
| (4) | High Street Fashion Store | Well-known high street fashion chain | Stock theft | Stock loss figures at audit | Y |
| (5) | Hereford Cathedral Toilets | Single unisex and disable toilet | Heroin-taking | Incidents logged by the Verger | Y |
| (6) | Aylestone Park | A children's playground, playing fields, recreation area | Speeding, and anti-social driving | Calls to the police | Y |
| (7) | Rural road-side land adjacent to a primary school | An area frequently misused for fly-tipping | Fly-tipping | Incidents reported to the police or noted by the school caretaker | Y |
| (8) | Huntsham Bridge area of the Wye Valley and Forest of Dean | Private forested estate and game area | Poaching | Poaching rates reported by gamekeeper | N |
| (9) | Hereford City Centre | A pedestrianised intersection | Cycling | Number of cyclists failing to dismount | Y |

### 5.2.3. Stimuli

The stimuli used were commercially available signs designed, purchased, and installed by Hereford Police. Figure 1 shows the priming signs used. Additional informational signs were also placed in some locations as indicated in Table 1 and these are shown in figure 2.

*Figure 5.1. Priming Signs Used in The Hereford Study*

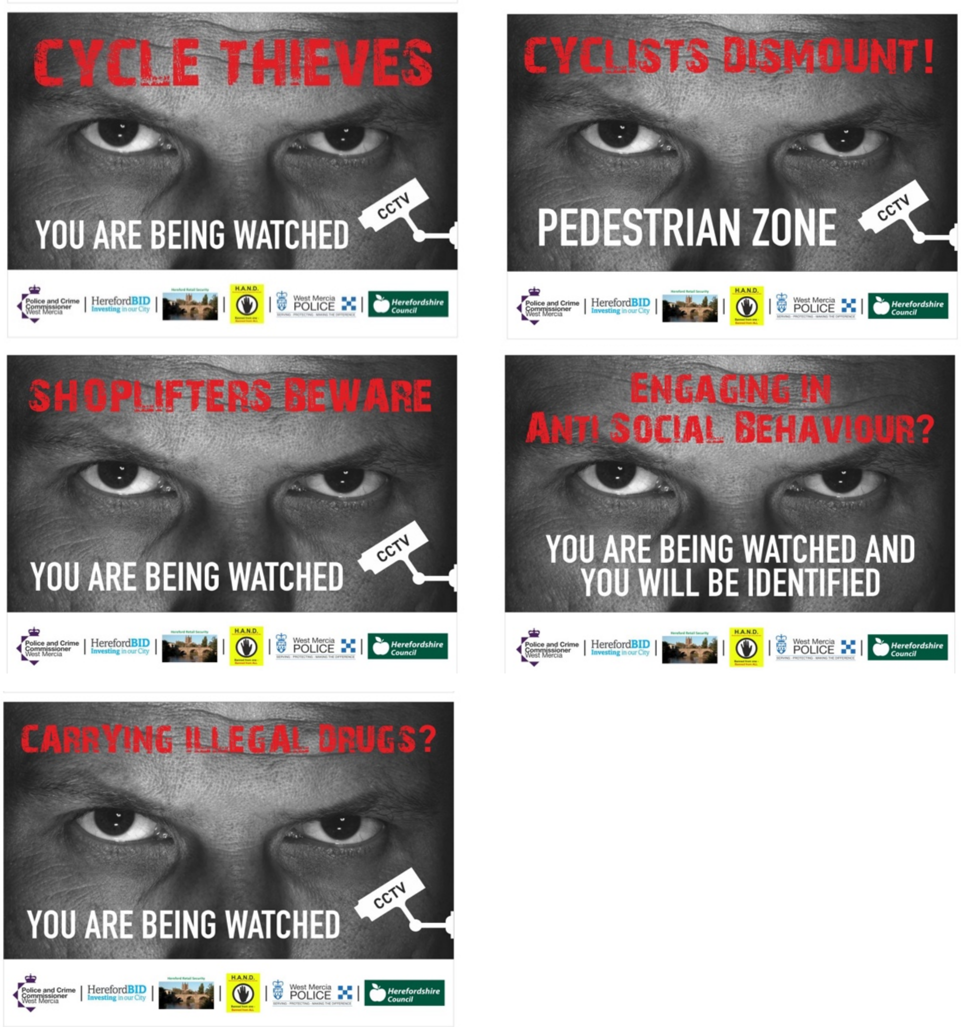
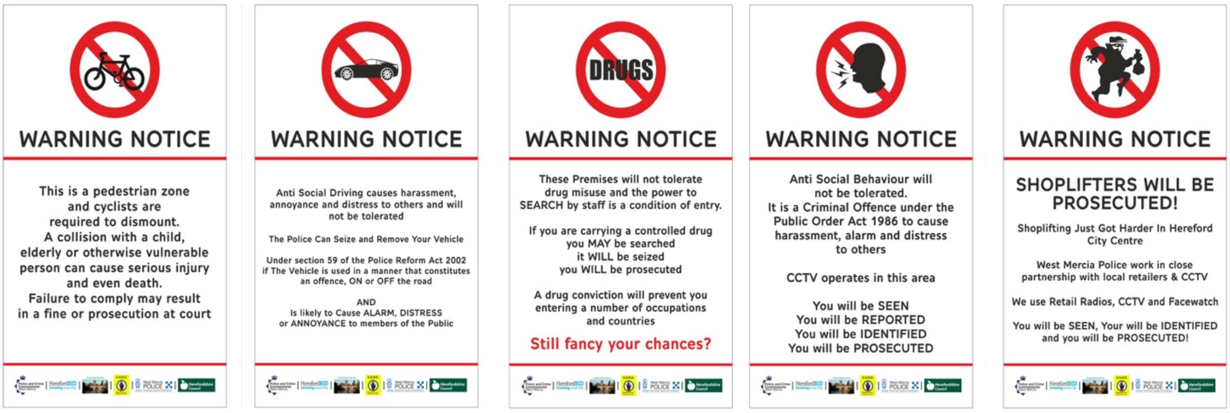

CYCLE THIEVES
YOU ARE BEING WATCHED
CCTV
CYCLISTS DISMOUNT!
PEDESTRIAN ZONE
CCTV
SHOPLIFTERS BEWARE
YOU ARE BEING WATCHED
CCTV
ENGAGING IN ANTI SOCIAL BEHAVIOUR?
YOU ARE BEING WATCHED AND YOU WILL BE IDENTIFIED
CARRYING ILLEGAL DRUGS?
YOU ARE BEING WATCHED
CCTV
Police and Crime Commissioner West Mercia
HerefordBID Investing in our City
H.A.N.D.
West Mercia POLICE
Herefordshire Council

*Figure 5.2. Additional Signs Used in Hereford*


WARNING NOTICE
This is a pedestrian zone and cyclists are required to dismount. A collision with a child, elderly or otherwise vulnerable person can cause serious injury and even death. Failure to comply may result in a fine or prosecution at court
WARNING NOTICE
Anti Social Driving causes harassment, annoyance and distress to others and will not be tolerated
The Police Can Seize and Remove Your Vehicle
Under section 59 of the Police Reform Act 2002 if The Vehicle is used in a manner that constitutes an offence, ON or OFF the road
AND
Is likely to Cause ALARM, DISTRESS or ANNOYANCE to members of the Public
DRUGS
WARNING NOTICE
These Premises will not tolerate drug misuse and the power to SEARCH by staff is a condition of entry.
If you are carrying a controlled drug you MAY be searched it WILL be seized you WILL be prosecuted
A drug conviction will prevent you entering a number of occupations and countries
Still fancy your chances?
WARNING NOTICE
Anti Social Behaviour will not be tolerated. It is a Criminal Offence under the Public Order Act 1986 to cause harassment, alarm and distress to others
CCTV operates in this area
You will be SEEN
You will be REPORTED
You will be IDENTIFIED
You will be PROSECUTED
WARNING NOTICE
SHOPLIFTERS WILL BE PROSECUTED!
Shoplifting Just Got Harder In Hereford City Centre
West Mercia Police work in close partnership with local retailers & CCTV
We use Retail Radios, CCTV and Facewatch
You will be SEEN, Your will be IDENTIFIED and you will be PROSECUTED!

### 5.2.4. Data Recording

Data recording was conducted by Hereford Police based on West Mercia Police's collation of calls made to the police seven months prior to the intervention and for the seven months of the intervention. Data was also provided by managers, staff and other personnel who were on site at each location. Additional qualitative data was collected via a survey designed as a simple questionnaire sent to business owners and data collectors by West Mercia Police asking for all quantitative data that might be relevant to be provided and for their qualitative assessment of the effectiveness of the intervention.

Both the police and I as the researcher requested detailed and carefully recorded data on shop lifting from the two retail outlets – the fashion store and the country clothing & gun store. In the case of the latter, despite prior agreement, data was only provided anecdotally with the store owner reporting that stock loss was little changed but that he'd found fewer discarded tags in the changing rooms after putting up the signs – discarded tags in the changing rooms being a sign of shoplifting. Clearly, we could not conduct analysis on this. At Hereford Cathedral the arrest of a known drug user during the intervention's data-recording period, in addition to an increased police presence in the area, confounded our results. Moreover, there were data recording limitations at the Cathedral that prevented reliable statistical analysis being undertaken. Specifically, the verger's records were at times unclear being maintained in a paper log, and the one post-arrest drug incident recorded turned out to be a Halloween prank. Likewise, in the intervention in the Forest of Wye to reduce poaching, data was not recorded with the rigour required to support statistical analysis. Finally, the intervention in the City Centre to encourage cyclists to alight when traversing a pedestrian area was ended prematurely by the local council, with insufficient data recorded to enable analysis. Aware that we might be accused of excluding negative findings it should be noted that at all those locations excluded, those involved suggested the intervention had been effective. Had we included their 'data' it would have strengthened our findings. See discussion for further details.

### 5.2.5. Data Plan.

Analysis of before and after data, that is, calls to the police in the case of McDonald's and Aylestone Park, cycle thefts as recorded by staff at the Hospital, and stock-loss figures in the case of the clothing store, were used to calculate odds ratios and to conduct the Pearson's Chi Squared based on expected values in lieu of a control condition. For count data with a value of zero the Haldane-Anscombe (H-A) correction was applied adding 0.5 to all count data (Lawson, 2004; Ruxton & Neuhäuser, 2012).

## 5.3. Results

Table 5.2. displays the results from our layered priming intervention. There were significant reductions in cycle theft at the hospital, $X^2(1) = 9.59$, *OR* = 0.45, *CI*= 0.00, 0.35, *p*<.01, on shop lifting at the High Street Fashion Store, $X^2(1) = 16.45$, *OR* = 1.38, *CI*= 1.11, 1.72, 2.32, *p*<.001, and on fly-tipping site at the rural roadside location, $X^2(1) = 29.03$, *OR* =0.02 , *CI*= 0.00, 0.14, *p*<.0001. There was no significant effect on antisocial behaviour at McDonald's, $X^2(1) = 10.08$, *OR* = 0.86, *CI*= 0.72, 0.96, *p* = .55, or on antisocial driving at Aylestone Park, $X^2(1) = 0.08$, *OR* = 0.85, *CI*= 0.42, 1.00, *p* = 0.78. Notably, there were reductions in crime or antisocial behaviour at every site where data was available. Problems with data collection arose at the other locations and no data was available for analysis. The data collection issues are described in the discussion.

Table 5.2.

**Frequency of Antisocial Behaviour at 9 Experimental Locations.** Table displays observed frequencies of antisocial behavior ($O_i$), before and after the intervention, and expected frequencies ($E_i$), along with the Pearson Chi Square statistic and associated statistics. *H-A* = indicates the Haldane-Anscombe correction was applied to counts of zero to enable calculation of expected frequencies. N/A = Not Applicable, those sites where data collection problems prevented evaluation.

| | | Anti-Social Behaviour Before | | Anti-Social Behaviour After | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Location | | $O_i$ | $E_i$ | $O_i$ | $E_i$ | $X^2(1)$ | *OR* | *CI* | *p* | |
| McDonalds | Anti-social behaviour | 38 | 35.5 | 33 | 35.5 | 0.35 | 0.86 | 0.72, 0.96 | 0.55 | |
| Hereford Country Hospital | Cycle theft | 11.5 | 6 | 0.5 | 6 | 9.59 | 0.45 | 0.00, 0.35 | <.01 | *H-A* applied |
| Country Clothing & Guns Store | Stock theft | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | |
| High Street Fashion Store | Stock theft | 374 | 322.5 | 271 | 322.5 | 16.45 | 0.72 | 1.11, 1.72 | <.0001 | |
| Hereford Cathedral Toilets | Heroin-taking | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | |
| Aylestone Park | Speeding, and anti-social driving | 7 | 6.5 | 6 | 6.5 | 0.08 | 0.85 | 0.42, 1.00 | 0.78 | |
| Rural road-side land adjacent to a primary school | Fly-tipping | 30.5 | 15.5 | 0.5 | 15.5 | 29.03 | 0.02 | 0.00, 0.14 | <.0001 | *H-A* applied |
| Huntsham Bridge area of the Wye Valley and Forest of Dean | Poaching | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | |
| Hereford City Centre | Cycling | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | |

## 5.4. Discussion

The franchise manager in McDonalds reported a dramatic reduction in anti-social behaviour in the restaurant following the start of the intervention. He suggested this was reflected in a c.70% reduction in the number of customer complaints recorded after the intervention began, but despite multiple enquires and requests, he was unfortunately unable to provide this data to the police or the researchers due to data storage and sharing policies within McDonalds – they do not hold data for longer than 6 months and don't make customer complaints data available externally. We found no effect of eye cues at McDonalds based on the number of calls made to the police before and after. Similarly, the Aylestone Park signage intervention was rated a success by the Park Association who reported that prior to the signs being erected there were incidents of anti-social driving in the park 5 nights out of 7, with this falling to almost zero after the signs were put up. However, this reporting was not considered sufficiently independent or systematic to be relied upon. Police calls data showed no significant before/after difference.

Our High Street Fashion store reported a significant reduction in the stock lost during the months the watching eyes intervention ran for, with losses down 27% from £374 per week to £271 per week following the erection of eyes signs in the changing rooms. Our statistical analysis supported their claim to significance $X^2(1) = 16.45$ $p < .0001$. Similarly, Hereford County Hospital reported a dramatic reduction in cycle theft, with 11 stolen in the 7 months prior to the intervention starting and 0 in the 7 months after the signs were put up $X^2(1) = 9.59$ $p < .01$.

The police received zero reports of fly-tipping post intervention compared with 30 in the 7 months prior, $X^2(1) = 29.03$ $p < .0001$. This result was so surprising that further enquiries were made to understand if anything else might have changed to bring about such a dramatic result. Both the police officer responsible for the area and the caretaker at the local school provided assurances that there had been nothing else to account for the change.

The Gamekeeper from Huntsham Bridge reported an 80% reduction in poaching after he placed the signs up, noting, without a hint of irony: “We must be getting under somebody’s skin because the signs were stolen!”. That the signs had been put up at significant height, thus requiring the use of a ladder for the removal, would seem to support his claim. However, the only reporting available was the Gamekeeper’s anecdotal evidence. No independently or systematically collected data was available to confirm his claims.

The present research replicates the dramatic effect of signs bearing watching eyes and text found by Nettle et al. (2012) whose study, if you recall, showed that cycle crime reduced by 62% over 12 months at sheds with the signs, and rose by 65% at sheds where the signs were not put up. In fact, our study went considerably further finding that the signs reduced cycle crime by 100%. Unfortunately, we were not able to track cycle crime at alternate sheds where the signs were *not* put up and thus were unable to either provide a baseline or to look for evidence of displacement. Similarly, dramatic results were found with the ending of fly-tipping following erection of the signs at a site which had seen 30 fly-tipping incidents previously, and saw zero after - a dramatic result that exceeded our expectations. These findings certainly suggest signage interventions can be effective, and suggest that at in this instance, watching eyes did not reduce the effectiveness of signs and may enhance their influence.

However, they may have done in our shoplifting intervention. We replicated the effectiveness of a simple signage intervention in reducing shoplifting, but our 27% reduction was not as great as the 91% reduction achieved by McNees et al. (1976) in their intervention in a Tennessee department store. Perhaps the text *is* key: in McNees et al.’s study their signs were text-only and displayed the messages: ‘Shoplifting is Stealing’, ‘Shoplifting is a Crime’, ‘Shoplifting is not Uplifting’, ‘Shoplifting Helps Inflation’. Consequently, we cannot discount the possibility that the eyes made the signs less effective than simple text signs.

Judging on police calls data alone the signs placed to reduce anti-social driving at Aylestone Park were not effective. Hereford Police believe this might be due to those in the local Neighbourhood Watch reporting more anti-social driving incidents after seeing the police put the signs up. We describe this limitation in the wider context of our studies below, but report the result to avoid accusations that we are selectively interpreting only the findings that conform to our hypothesis.

Recent research offers an intriguing possible additional moderator of the ‘watching eyes’ effect that may have influenced our results. Zuo, Huang, Cai, and Wang (2018) found that residential mobility moderated the effect of surveillance cues on anti-social behaviour, which they measured based on a participant’s willingness to lie in a test to gain a reward. Where those that moved often generally engaged in more anti-social behaviour, the presence of a surveillance cue correlated with a reduction in their anti-social behaviour. In contrast, those with low residential mobility increased their anti-social behaviour when a surveillance cue was presented. From this, we might expect eye cues to be more effective in reducing crime in areas with high social mobility and a large transitory population, such as a university campus or a hospital but to be ineffective in areas where there is low-social mobility and a small transitory population, such as villages and small towns. Hereford is the largest town in the county of Herefordshire with a population of 55,000 and is a regional tourist destination, so the results in the department store and the hospital might be due to there being a high transitory population. It seems probable that fly-tipping in an area is not a crime carried out by those who live there – who would dump rubbish on their own doorstep? So again, Zuo, Huang, Cai, and Wang’s (2018) finding may partly explain this result. It does not, however, shed any light on the measured ineffectiveness of the signs in discouraging anti-social driving at Aylestone Park – since Hereford Police advise that it is local youths that are believed to be largely responsible for this.

## 5.5. Limitations

The compromises required in running a real-world field study in collaboration with law enforcement authorities who view the study as an intervention to reduce crime are apparent throughout the design and prevent us isolating the effects of the various cues used.

Additional potential confounds arise from the reporting mechanism. We had to obtain consent for the signs to be deployed from the responsible guardian at each location. This took place immediately before the signs were put up, at which point the police asked for all relevant incidents to be reported so that an accurate record could be made of incidents. The ‘before’ data was taken from police databases, and, guardians admitted, during this period incidents had gone unreported as there was a perception that the police couldn’t, or wouldn’t, be able to attend and deal effectively with the issue. Consequently, it is likely that the data for the ‘before’ period under-accounts for the number of incidents, while the data for the ‘after’ period, that of the intervention, over-accounts in comparison. This limitation was considered in advance but it was not possible to put signs up without the permission of those whose property they were to go on. If anything, though, this would make it more likely we would have found non-significant effects, with the under-counting of crimes occurring before the intervention began.

There are also inherent limitations in before/after data given the possibility that any change in what is measured might be due to wider environmental fluctuations. A seven-month measurement period also risks climatic or seasonal influences confounding results. Given more time it would be preferable to compare a full twelve-month before/after period so that an appropriate baseline can be determined.

Our findings could have been considerably strengthened with better data collection, thus our results risk underplaying the effectiveness of the interventions. At the country clothing store it was reported that shop lifting was down post-intervention. At the Cathedral there were no post-incident drug-taking incidents compared with ten incidents before. Our gamekeeper reported an 80% reduction in poaching post-intervention. We are cautious in the context of the wider research in this thesis of extrapolating too much from our findings, but there is an attendant risk of under-playing our results here which at the very least urge further research given some of the dramatic reductions in crime that correlated with the deployment of our eyes and text signs.

## 5.6. Further Research

Further research must more carefully balance the interventions, isolating the test variables. It must test over a full twelve-month period, select a comparable control area in which to measure anti-social behaviour over the same period before and after the intervention. It would also be desirable to measure over a wider area concurrently in order to test for displacement or diffusion effects. Tracking data closer to source, as we did with the department store's stock loss and cycle crime at the hospital, rather than using phone calls to the police as a proxy, would also provide more reliable insight

# 5. Study 5. Watching Eyes on Crime in Hereford: A Collaborative Field Experiment

# Chapter 6

# 6

# Study 6 “Cycle Thieves: We’re Watching You”. A Collaborative Field Experiment and Replication Attempt.

> *“Surveillance breeds conformity.”*
>
> Glenn Greenwald, 2014.
>
> American lawyer & journalist whose reports helped Edward
> Snowden reveal NSA & GCHQ surveillance in June 2013.

In this chapter we explore the effect of perceived surveillance on behaviour in the field in order to replicate and extend Nettle, Nott and Bateson’s influential (2012) study of the effect of watching eyes on cycle crime. The study took place in the London Borough of Ealing. As with Nettle et al., the study was conducted over a 12-month period. The signs used were as similar as possible to those used by Nettle et al, a condition we describe as ‘eyes and text’ to reflect the sign’s design. To allow us to separate out the effects of different cues we extended Nettle et al.’s design by including signs with just images of eyes and signs with just images of text. We compared before and after cycle theft data provided by the Metropolitan Police in our eyes & text, eyes only, and text only conditions across a total of 14 sites, with before and after data against a baseline control of cycle theft Borough-wide (excluding our study sites). No significant effect of any of the interventions: eyes and text, eyes only, or text only, on the frequency of cycle theft was found.

# 6. “Cycle Thieves: We’re Watching You!” A Collaborative Field Experiment & Replication.

## 6.1. Introduction

In our final study we attempted to replicate Nettle, Nott and Bateson’s influential (2012) study of the effect of signs with eyes, the words ‘Cycle Thieves: We Are Watching You’ a police logo and ‘Operation Crackdown’, on cycle theft (Figure 6.1). Nettle et al. report that erecting the signs correlated with a 62% decrease in cycle crime as measured over the course of a 12 month intervention period at 3 cycle sheds when compared with theft at the same sheds over a 12 month ‘before’ control period. Over the same period control locations – comprised of all other cycle sheds across Newcastle University’s campus - saw an increase in cycle theft of 65%. Nettle et al. suggest this may have been the result of cycle theft being displaced by the signs.

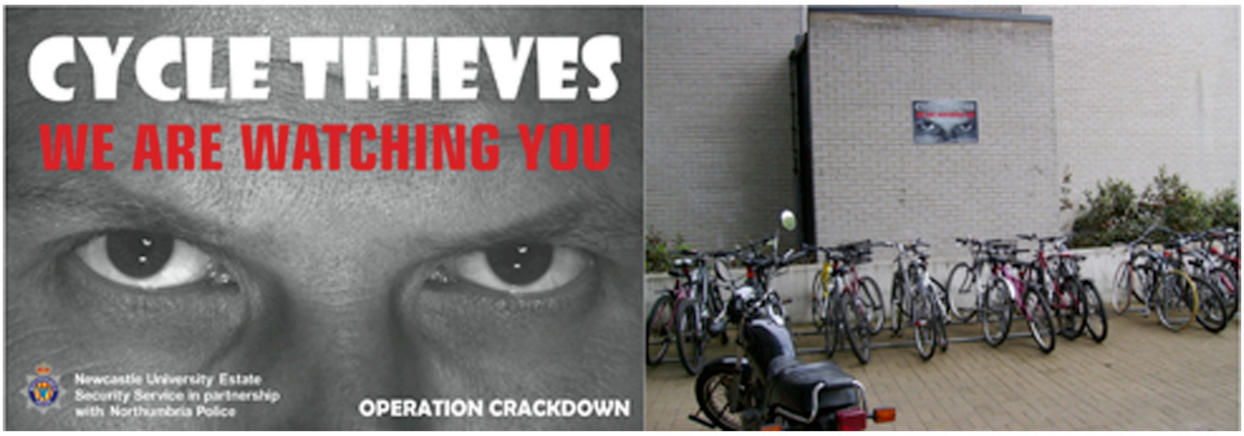

CYCLE THIEVES
WE ARE WATCHING YOU
Newcastle University Estate Security Service in partnership with Northumbria Police
OPERATION CRACKDOWN

*Figure 6.1. Sign as used by Nettle et al. (2012), and the sign in location at one of the cycle sheds at Newcastle University’s Campus.*

Reviewers of our meta-analytic findings have questioned whether Nettle et al.’s cycle crime experiment ought to be considered a watching eyes study. They objected to the studies use of the text and police logos, which might have been sufficient to reduce crime irrespective of the presence of eyes. Nettle et al. acknowledge this limitation in their report of their findings stating that “*our current design did not separate out the effects of merely installing any sign at all from the contents of these signs, or, within the contents of these signs, separate out the effects of the verbal and the image components*”. Our study aimed not just to replicate Nettle et al.’s findings, but to extend their paradigm to separate out the effects of eyes and text.

## 6.2. Method

The study was organised and coordinated with Ealing Council and the London Metropolitan Police Service.

### 6.2.1. Design

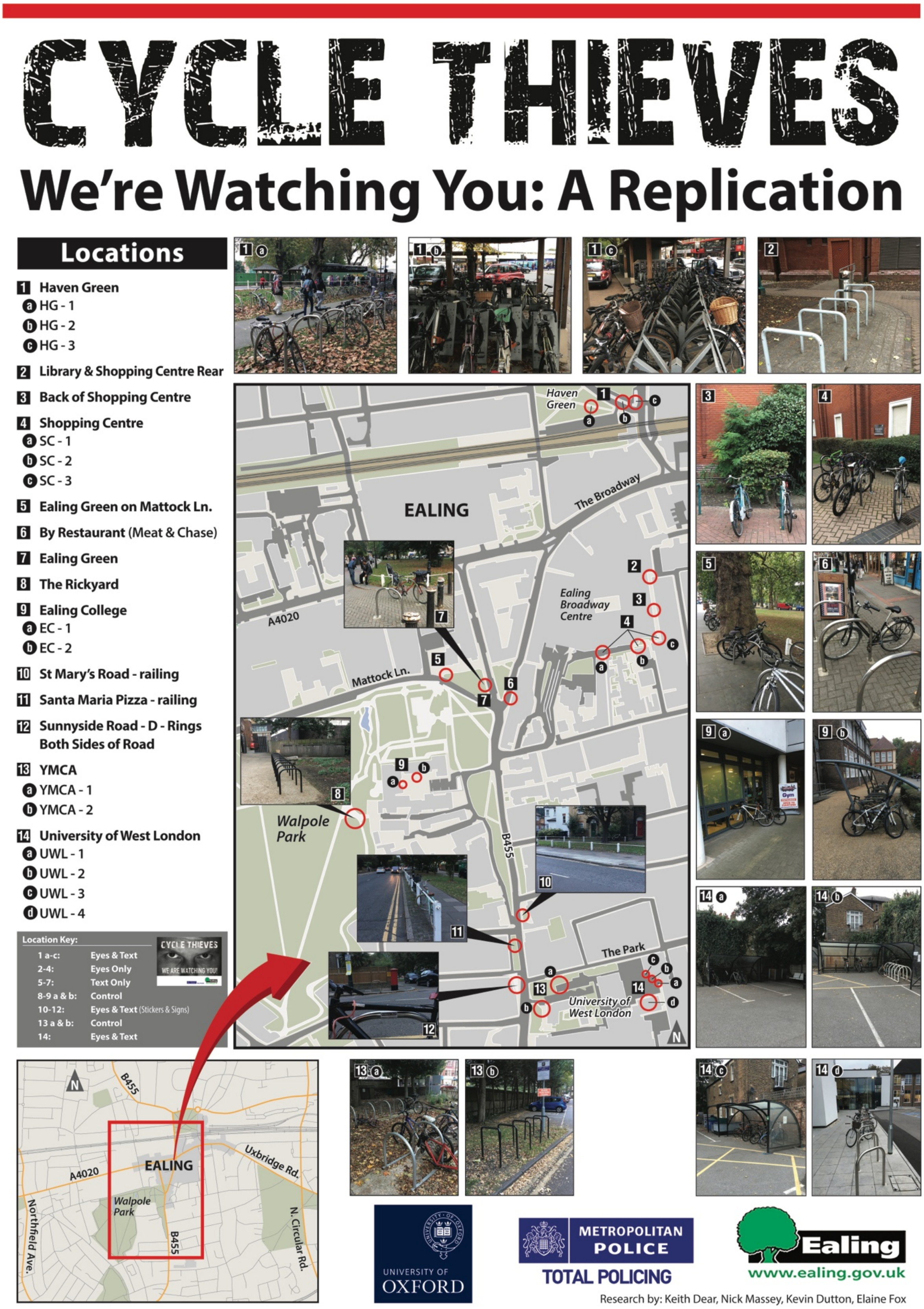
Sites were selected in consultation with Ealing Council. The locations were cycle sheds in what were judged by the Council to be high cycle crime areas. Cycle theft figures for precise locations were not available at the time of site selection.

### 6.2.2. Study Sites

Study sites are described in the infographic at Figure 6.2. Control locations were initially planned to be the specific locations described in the infographic but this was later widened to include Borough-wide cycle theft data, excluding the experimental locations, due to low count data across the specified control locations both before and after the intervention.

6. Study 6 “Cycle Thieves: We’re Watching You!” A Field Experiment & Replication.

*Figure 6.2. Ealing Cycle Crime Infographic*



# CYCLE THIEVES

## We’re Watching You: A Replication

### Locations

- **1 Haven Green**
  - ⓐ HG - 1
  - ⓑ HG - 2
  - ⓒ HG - 3
- **2 Library & Shopping Centre Rear**
- **3 Back of Shopping Centre**
- **4 Shopping Centre**
  - ⓐ SC - 1
  - ⓑ SC - 2
  - ⓒ SC - 3
- **5 Ealing Green on Mattock Ln.**
- **6 By Restaurant** (Meat & Chase)
- **7 Ealing Green**
- **8 The Rickyard**
- **9 Ealing College**
  - ⓐ EC - 1
  - ⓑ EC - 2
- **10 St Mary’s Road - railing**
- **11 Santa Maria Pizza - railing**
- **12 Sunnyside Road - D - Rings Both Sides of Road**
- **13 YMCA**
  - ⓐ YMCA - 1
  - ⓑ YMCA - 2
- **14 University of West London**
  - ⓐ UWL - 1
  - ⓑ UWL - 2
  - ⓒ UWL - 3
  - ⓓ UWL - 4

**Location Key:**

| | |
|---|---|
| 1 a-c: | Eyes & Text |
| 2-4: | Eyes Only |
| 5-7: | Text Only |
| 8-9 a & b: | Control |
| 10-12: | Eyes & Text (Stickers & Signs) |
| 13 a & b: | Control |
| 14: | Eyes & Text |

CYCLE THIEVES
WE ARE WATCHING YOU!
1a
1b
1c
2
3
4
5
6
7
8
9a
9b
10
11
12
13a
13b
14a
14b
14c
14d
Haven Green
EALING
The Broadway
Ealing Broadway Centre
A4020
Mattock Ln.
Walpole Park
B455
The Park
University of West London
N
B455
EALING
A4020
Uxbridge Rd.
Walpole Park
Northfield Ave.
N. Circular Rd.
B455

UNIVERSITY OF OXFORD

METROPOLITAN POLICE
TOTAL POLICING

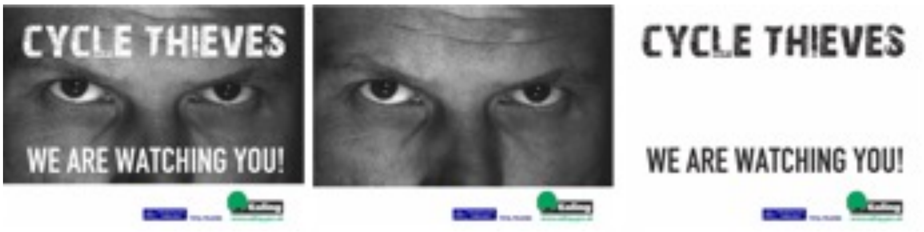
Ealing
www.ealing.gov.uk

Research by: Keith Dear, Nick Massey, Kevin Dutton, Elaine Fox

### 6.2.3. Stimuli

Signs were put in place by the council over a period of fourteen days in late February 2017. The first sign aimed to mimic Nettle et al.’s, the eyes only sign omitted the text and the text only sign omitted the eyes.

*Figure 6.3. Signs as used in the Ealing Study*


CYCLE THIEVES
WE ARE WATCHING YOU!
CYCLE THIEVES
WE ARE WATCHING YOU!

### 6.2.4. Data Recording

Data from the Metropolitan Police was provided for the 12 month ‘before’ period from 01/02/16 to 31/01/17 and for the 12 month ‘after’ period 01/02/17 to 31/01/18.

The data was only available after the intervention had commenced due to delays in the processing of our data release request.

### 6.2.5. Data Plan

Our methods mirrored those of Nettle et al. for ease of comparison. We used Fischer’s Exact Test to look for associations between intervention and the number of thefts. We followed Nettle et al. in expressing our effect size as an odds ratio.

To investigate any variance in the effect of the intervention over time we followed Nettle et al. in splitting our intervention period into two 6-month blocks, calculating odds ratios for each.

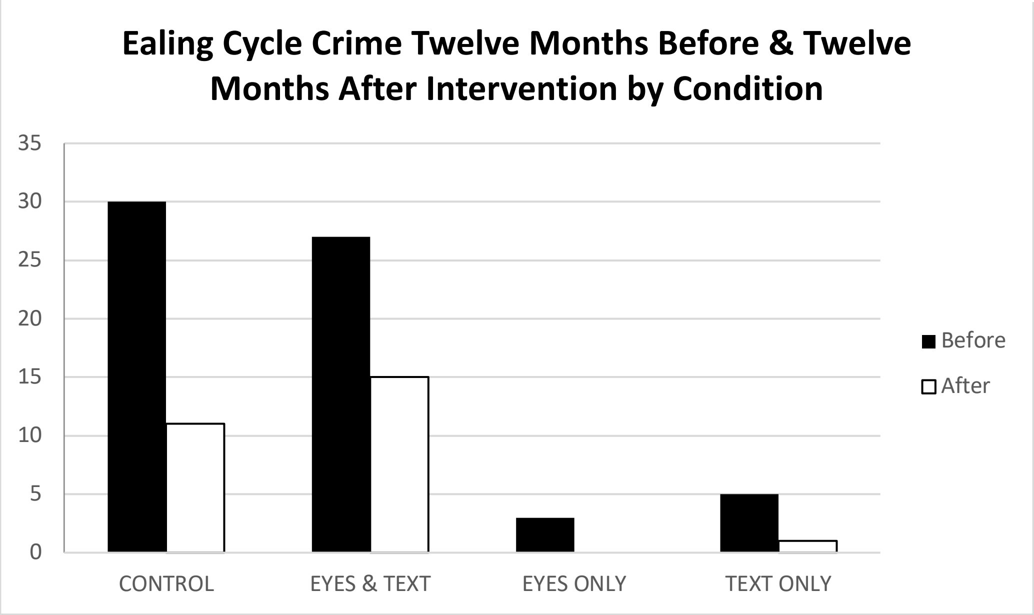
## 6.3. Results

There was a 63% reduction in cycle theft across the control locations (all areas of Ealing outside of our intervention sites). This compares with the smaller reduction of 41% at our Eyes & Text condition resulting in the non-significant findings. Although there were greater reductions in cycle theft at our eyes (100%) and text location (80%) when compared with the control (63%), the low frequency of cycle theft in these conditions resulted in the non-significant results and wide confidence intervals. Aggregating across all of our intervention sites and comparing them with the baseline control also yielded non-significant results. Table 6.1 and Figure 6.4. summarise cycle thefts by condition before and after the intervention.

Table 6.1. Ealing cycle theft: comparing 12 months before with 12 months after.

| | | | | Fischer Test | | |
|---|---|---|---|---|---|---|
| Condition | Before | After | % Reduction Before/After | *OR* | *CI* | *P* |
| Control | 30 | 11 | 63% | | | |
| Eyes & Text | 27 | 15 | 41% | 1.51 | 0.54, 4.32 | 0.48 |
| Eyes Only | 3 | 0 | 100% | 0 | 0.00, 7.46 | 0.56 |
| Text Only | 5 | 1 | 80% | 0.55 | 0.11, 5.79 | 0.55 |
| Total for Intervention Sites | 35 | 16 | 58% | 1.24 | 0.46, 3.46 | 0.65 |

*Figure 6.4. Ealing Cycle Crime Bar Chart by 6 Monthly Period Over 12 Months Before & After*



Comparing the data by six-monthly periods to see whether there were any time-variant effects (after Nettle et al., 2012) we see (Tables 6.2. and 6.3., Figure 6.5.) that there were no cycle thefts in the first six months after our intervention in any condition, including the control. Consequently, we could not compute odds ratios for this data. In the second six months cycle crime resumed in Ealing. There were no statistically significant differences between our intervention conditions and control. The 100% reduction in cycle theft comparing the second six months before the intervention with the second six months prior was based on just 3 thefts before compared with 0 after and did not reach statistical significance, again because of the low frequencies of cycle theft.

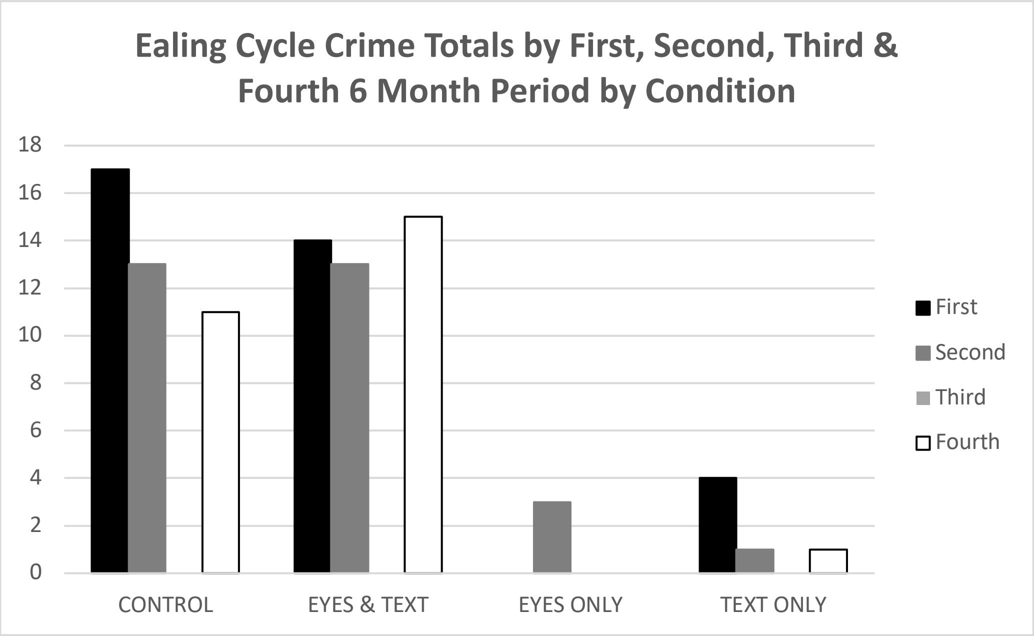
Table 6.2. Ealing Cycle Thefts: First vs Third 6 monthly period

| | Before (No Signs) | | After (Signs Erected) | | | Fischer Test | | |
|---|---|---|---|---|---|---|---|---|
| Condition | First | Second | Third | Fourth | % Reduction first 6 months before vs first 6 months after | *OR* | *CI* | *P* |
| Control | 17 | 13 | 0 | 11 | 100% | | | |
| Eyes & Text | 14 | 13 | 0 | 15 | 100% | - | - | - |
| Eyes Only | 0 | 3 | 0 | 0 | - | - | - | - |
| Text Only | 4 | 1 | 0 | 1 | 24% | - | - | - |

Table 6.3. Cycle Thefts: Second vs Fourth 6 monthly period

| | Before (No Signs) | | After (Signs Erected) | | | Fischer Test | | |
|---|---|---|---|---|---|---|---|---|
| Condition | First | Second | Third | Fourth | % Reduction second 6 months before vs second 6 months after | *OR* | *CI* | *P* |
| Control | 17 | 13 | 0 | 11 | 15% | | | |
| Eyes & Text | 14 | 13 | 0 | 15 | +15% (increase) | 1.36 | 0.40, 4.68 | 0.78 |
| Eyes Only | 0 | 3 | 0 | 0 | 100% | 0 | 0.00, 3.45 | 0.24 |
| Text Only | 4 | 1 | 0 | 1 | - | 1.17 | 0.01, 99.73 | 1 |

*Figure 6.5. Ealing Cycle Crime Bar Chart by 12 Monthly Period*



## 6.4. Discussion

Although our experiment has limitations it did not replicate Nettle et al.’s (2013) findings. Given there were no thefts across Ealing for 6 months after our signs were erected we might appeal to diffusion effects as an explanation – suggesting that emplacing the signs brought benefits over a wider area – and cautiously infer that placing anti-cycle crime signs can reduce cycle crime perhaps irrespective of the nuances in design, but that these effects are time limited. We could with even greater caution suggest that eyes signs may be more effective than eyes and text or text only signs, given the 100% drop in crime over 12 months at the eyes signs sites. However, as will be clear from the limitations, further research is needed.

Our data is problematic. The absence of any cycle crime in the six months after the signs were put up could be a dramatic demonstration of the effectiveness of erecting anti-cycle crime signs irrespective of design. To explain: our experimental signs were distributed across the Borough so it is possible that diffusion or displacement effects could account for the drop in cycle crime in the control areas; diffusion, with signs having effects beyond their immediate locale; displacement with cycle crime pushed to neighbouring boroughs. However, it is also possible that the absence of cycle crime over this six-month period is a simple coincident anomaly.

The evidence suggests that in the second six months after the signs were emplaced the eyes and text signs were not effective. Low counts in the separate text and eyes conditions makes comparison unreliable and statistical analysis too low in power to detect significant differences. But thefts in the ‘eyes & text’ condition increased compared with the respective 6 months before the intervention. This occurred despite a reduction in cycle crime across Ealing. Thus, we did not replicate Nettle et al.’s effects.

## 6.5. Limitations

The low frequency of cycle crime in all conditions is the major limitation of our study. Intervention sites were selected by Ealing Council based on reported cycle crime ‘hotspots’ as proposed by Ealing’s Safer Communities Analysis and Research Team. These were broad areas. We did not have detailed locational cycle theft data from the London Metropolitan Police until after the study had begun. Consequently, the specific locations chosen within the broad hotspots were selected by Ealing Council based on the permissions of the landowners, the presence of cycle sheds or ‘D rings’ – the bars in the ground to which bikes can be locked – and local knowledge of areas perceived to be high in cycle crime. The data presented shows that a number of the sites selected did not have sufficient cycle thefts to enable statistically significant conclusions to be drawn regarding the effectiveness of the interventions.

Additionally, Ealing Council had a major concern with ‘street clutter’ – the tendency of signs to proliferate and look ugly. The consequence of this was that the cycle shed at Haven Green, the largest cycle shed with the highest cycle theft rate, which we included in our ‘Eyes & Text’ condition, had just two signs that were less prominent that would have been desirable.

# Chapter 7

# 7

# Discussion

> *"We cannot perceive the world and at the same time apprehend a look fastened upon us; it must be either one or the other."*
>
> Jean-Paul Sartre (1978, 258)

## 7. Discussion

### 7.1. Overview

This final chapter is divided into three sections. First, it provides an updated meta-analytic effect size from that reported in Chapter 3 by adding in the three Joy of Destruction studies and two field experiments on eye cues and antisocial behaviour reported in chapters 4,5, & 6. Second, it summarises and discusses the findings reported throughout the thesis. Third, it highlights some limitations of both the studies reported in Chapters 2, 4, 5, and 6 and the wider literature, before suggesting directions for future research.

## 7.2. Updated Meta-Analysis of the Effect of Eye Cues on Antisocial Behaviour

In an updated meta-analysis, we amalgamated the counts of antisocial behaviour in Hereford for before and after periods across all the interventions where data was collected in Study 5 using the expected frequencies for the before and after period in the control. The Haldane-Anscombe correction for zero counts was not applied as this was considered to be unnecessary for the aggregated data. Similarly, for Study 6, we amalgamated the cycle theft counts for the locations with the eyes signs with those with the eyes and text signs, comparing this data against the counts of cycle theft in the control locations before and after the intervention period.

Using the procedure described in chapter 3, the fixed effects model including the additional 5 studies returned a result of $OR_{log} = -.15$, *SE* = .04, 95% *CI* [-.23, -.08], $p <$ .0001); exponentiating the $OR_{log}$ we get an *OR* = 0.86, suggesting a 14% reduction in ASB in the eyes when compared with the control condition. The $Q$-test for homogeneity of variance was significant ($Q(19) = 81.76$, $p < .0001$) again requiring us to fit a random effects model to the data.

The random effects model suggested that there was a 29% reduction in the risk of antisocial behaviour between the eyes and control condition ($OR_{log} = -.34$, *SE* = .12, 95% *CI* [-.56, -.11]) $p < .01$; *OR* = .71). See Table C1 in Appendix C at the end of this thesis for an updated meta-analytic table.

As may be seen, combining the results of studies 3, 4, 5, and 6 does little to change the overall meta-analytic effect size. Chapter 3 found a 35% reduction in the risk of antisocial behaviour when watching eyes were present. The updated meta-analytic effect size reduces the estimate to a 29% reduction in antisocial behaviour in the presence of eye cues.

## 7.3. Summary of Chapters 1-6

The first part of Chapter 1 framed the research that has followed in the philosophical context of Bentham and Foucault's arguments that just feeling watched might be enough to change behaviour, perhaps leading to greater efficiency, rule adherence and social cohesion, or perhaps leading to self-censorship, or maybe both. It then outlined just how widespread surveillance has become in the 21st Century and demonstrated the range of domains in which Bentham and Foucault's arguments are still being advanced – from a claimed 'Ambient Protective Effect' of technologies of surveillance in peacekeeping to claims for 'chilling effects' leading internet users to self-censor online in the months after Edward Snowden released his revelations on the extent of State surveillance in the USA and UK. It sought to demonstrate the need for empirical evidence to help advance a debate on the effects of surveillance that began at least as early as Plato's writing in the 5th Century BCE and remains pressingly relevant today.

The second half of Chapter 1 provided an inter-disciplinary overview of surveillance research to provide immediate insight into the philosophical and practical questions raised in the first part of Chapter 1 and to frame and inform the research in Chapters 2-6. We first argued that psychological science is under-represented in the wider field of surveillance studies, noted how central understanding the effect of being watched remains for contemporaneous debate in psychology – with authors debating the extent to which the presence of cameras may or may not have altered outcomes in recent failed replication attempts. From there we identified research into the watching eyes effect as a promising route to understanding whether merely feeling watched – as Bentham and Foucault argued –

really might change behaviour unconsciously independent of explicit, conscious concerns regarding future consequences.

The final part of Chapter 1 presented the evolutionary, ethological, neuropsychological, and biological evidence - along with insights from computer science and research with brain injured patients - that suggests humans *are* hyper-sensitive to gaze and sometimes to false gaze cues. We saw how this hypersensitivity to being watched may have underpinned the evolution of cooperation – enabling us to rapidly ‘perspective take’ and see ourselves as others see us to adapt our behaviour, avoid censure, and enable cooperation. We saw how this hypersensitivity may have helped us to communicate pre-linguistically, and, that the shape and colouration of our eyes may have evolved to perform a communicative function, giving rise to a rich ‘language of the eyes’ - to which we pay attention rapidly and automatically. The evidence suggested that just feeling the gaze of others upon us or upon the objects we look at changes the neural pathways through which we process visual stimuli as well as changing the neural pathways through which we initiate and enact motor control. We saw how the evolutionary pressure of natural selection might have made a hypersensitivity to feeling watched adaptive in three principal ways. First, to enable us to spot prey to eat, and predators to avoid being eaten. Second, to ensure we conformed to norms to enable cooperation and reap the benefits of living in larger groups. Third to sustain our reproductive fitness within groups by avoiding behaving before the gaze of others in a way that might invite reputational damage, punishment and reductions in our status. We saw how cultural evolution might have extended our sensitivity to being watched to the point where we could feel observed by Gods and spirits – and how it is hypothesised that this allowed groups to expand beyond the limit of their ability to remember all the members of our communities.

# 7. Discussion

In support of this hypothesis we reviewed the neuropsychological evidence to suggest we experienced the presence of Gods, spirits and other agentive phenomena much as we experience being watched by others – with changes in the neural pathways by which information is processed, and some evidence for behavioural changes - increased prosociality and reduced norm-breaching - when cued to feel watched by this more abstract ‘supernatural surveillance’. Finally, we saw how psychology has explored surveillance effects since the 1960s, identified the useful distinction in surveillance research between private self-consciousness – where being made to be aware of ourselves cues us to try to live up to how we wish to see ourselves, and public self-consciousness where we seek to live up to what is expected of us by others – be it a specific audience or society. We noted that while there was evidence that the presence of an audience or the ‘mere presence’ of others could influence behaviour the evidence for the direction of any such effect was highly inconsistent and, consequently, the most crucial evidence – that which would determine how feeling watched changes behaviour, remained limited.

Chapter 2 reported the findings of Study 1, an investigation of whether the presence of eye cues, a camera (which wasn’t switched on and was ‘merely present’ – participants were *not* told it was there), or a camera which was switched on and which participants were told would be ‘filming but not able to see their actions in the task’, might increase helping behaviours in a simple computer based task – where helping was incidental to the task’s goal – or increase generosity in an economic game (the dictator game). No surveillance cue effect was found, nor was there any effect of the surveillance cues on self-monitoring – the extent to which one seeks to manage one’s reputation before the eyes of others. Additionally, although there was a main effect of surveillance cue on emotionality scores within the HEXACO PI-R personality model, the most likely explanation for this is that it was a consequence of a failure of our randomisation procedure.

# 7. Discussion

Chapter 3 reported the findings of a systematic review and meta-analysis of eye cues' effect on anti-social behaviour. We saw how gaze might prompt us to be more aware of our reputation, and thus looked to the evidence for the effect of given actions on reputation to better understand what effects we might expect eye cues, as a proxy for gaze, to cause. Relying principally on Klein and Epley's work (2014, 2015) we showed how the effects of generosity on reputation are not monotonic – while being fair rather than unfair improves reputation, greater generosity brings less benefit, and extreme generosity may not benefit reputation at all. In contrast, drawing principally on Baumeister's et al.'s (2001) review we saw that negative actions are consistently bad for a person's reputation. Thus, eye cues might have more consistent effects in reducing antisocial behaviour than in promoting prosocial behaviour. We saw how Northover et al. (2017a) had demonstrated that the effects of eye cues on generosity had been highly variable and inconsistent leading to a very small meta-analytic effect size while our own meta-analysis on eye cues effect on antisocial behaviour found a larger effect size that equated to a 35% reduction in the risk of anti-social behaviour when eye cues were present.

Chapter 4 used the Joy of Destruction task in Studies 2, 3 & 4 to test the hypothesis that eye cues – or the presence of a camera – might have more consistent effects in reducing antisocial behaviour. In the task, participants could pay to reduce another's income thereby hurting themselves to make others worse off – what we termed 'malicious behaviour'. Study 2 was a large-scale experiment under laboratory conditions conducted at Defence establishments across the UK and Study 3 was an online replication of the first. In both Study 2 and Study 3 there was a one-third chance that participants had some of their income 'destroyed' by chance in each round of the ten rounds in which they played the Joy of Destruction game. This was the case *regardless* of what their anonymous 'partner' in each round decided to do. Other researchers have described how this 'hides' the agentive act of destruction, and showed how removing this 'hidden' aspect reduced the frequency with which participants chose to destroy – perhaps because the psychological distance between participants was reduced when decisions were not hidden (Abbink & Herrmann, 2011; Abbink & Sadrieh, 2009). It is possible that while eye cues or a camera were cueing participants to feel more watched and *reducing* the perceived psychological distance between them and their 'partners', the role of chance was, simultaneously, offsetting this effect by *increasing* the perceived psychological distance. Consequently, in Study 4, we removed the role of chance. This made no difference to our results. Reviewing the evidence from the three Joy of Destruction studies (Studies 2, 3 & 4), no evidence was found for an effect of eye cues on malicious behaviour. Nor was there any effect of the presence of eye cues or a camera on public self-consciousness. Finally, there were no consistent interaction effects between surveillance cues and our exploratory individual difference measures. This applied to social value orientation [how prosocial a person was], a person's independent to interdependent self-construal [how much their identity merged with group identity], as well as their degree of religiosity.

Chapter 5 reported a large-scale field experiment (Study 5) conducted in Hereford in conjunction with West Mercia Police. Pictures of watching eyes along with various injunctive messages such as 'Cycle Thieves: We are Watching You!' were placed at a nine locations across the town of Hereford. Data was collected for seven months prior to the signs being erected, and seven months after, and included both police calls data, crime recorded by the police, and data collected by 'guardians' at each site – for example stock loss data at a major high street fashion retailer to measures shoplifting rates. Inadequate or unreliable data collection prevented us drawing conclusions from six of the nine locations – although anecdotal evidence suggested that all of these locations *had* seen a reduction in crime after the signs were erected. There were statistically significant reductions in crime [a 33% reduction in shop lifting at a major High Street fashion store where the signs had been placed in the changing rooms, a 100% reduction in cycle crime at cycle sheds at Hereford Hospital and a 100% reduction in fly-tipping at a roadside site] from locations in which data was reliable. Indeed, there was no site where the signs were erected and data was collected that didn't experience a reduction in crime in the seven months following the signs' installation. However, there was no baseline control area against which to measure our interventions, and thus the design could not control adequately for the possibility that reductions in the 'after' period were reflective not of the effectiveness of the intervention but rather of more general reductions in crime across the wider area.

Chapter 6 reported a large-scale attempt to replicate Nettle et al.'s (2012) study of the effects of eye cue signs on cycle theft at Newcastle. Nettle et al. had found that cycle crime reduced by 62% over 12 months after signs were erected at cycle sheds, and rose by 65% at sheds where the signs were not put up, measured against a 12-month period prior to the signs going up at both the intervention sites and the control sheds. One criticism of Nettle et al.'s study (and one which they themselves had noted) was that it did not separate the effect of just putting up a sign – which previous research in a retail shop had shown can reduce crime (McNees et al., 1976) - from the effect of eye cues, a criticism that might also be levelled at the Hereford study reported above. Consequently, Study 6's replication attempt conducted in the London Borough of Ealing, included three conditions: signs were erected that contained (1) just eye cues; (2) just the text 'Cycle Thieves We Are Watching You!' and (3) both eyes and text. A baseline was taken from cycle crime rates at non-intervention sites across Ealing and was recorded by the London Metropolitan Police Service over 12 months before the signs were put up and 12 months after. Unexpectedly, there was no cycle crime reported at all across Ealing in the 6 months after the signs were put up. However, from month 6 to 12 there were 11 thefts in areas with signs compared to 30 cycle thefts in the control locations representing a reduction of 63%. Breaking this down further, theft fell by proportionately more at the 'eyes only' sign locations: falling by 100%, and at the text only sites, falling by 80%, but the overall frequencies were so low (Eyes Before $N = 3$, Eyes After $N = 0$; Text before *N* = 5, Text After *N* = 1) that these reductions were not significant and are unreliable. The locations with ‘eyes and text’ signs, where counts were higher (Before *N* = 27, After *N* = 11) saw the smallest reduction in cycle theft of 41%. Our design was sound, but the late receipt of the ‘before’ data detailing frequencies of cycle crime meant that our intervention sites were not optimal. The unforeseeable absence of any cycle crime across the Borough of Ealing after our signs were put up might – at a stretch – be said to be the result of diffusion effects from our signs being erected. More likely, however, is that this was just an unrelated fluctuation in the crime rate, that coincidently confounds Study 6 and renders it, ultimately, uninformative.

## 7.4. Summary of Research Findings

Our consistent failure to find any significant increase in public self-consciousness in the presence of eye cues or a camera in Studies 1, 2, 3, or 4 might be explained simply by concluding that eye cues or the presence of a camera don’t make people feel more watched. If so, our failure to find any effect of surveillance cues on helping behaviours, generosity and malicious behaviour is not surprising. Set in the context of contemporaneous concerns regarding the wider reproducibility of the findings in psychological science (Aarts et al., 2015), and the recent meta-analyses on eye cues effect on moral judgement and generosity (Northover et al., 2017a, 2017b) this conclusion deserves to be given serious consideration.

However, Bradley et al.’s (2018) review of the effect of observability on prosociality did find a significant effect of *perceived* surveillance on prosociality in contrast to Northover et al.’s (2017a) findings. This was partly because their review included studies that primed perceived observability in other ways as well as with eye cues (e.g. with sentence completion tasks where the words were related to surveillance) and partly because their exclusion criteria allowed the inclusion of aggregated results from watching eyes studies such as Powell et al.’s

(2012) study on charitable donation at a supermarket that measured differences in donation per 1000 customers. In contrast, Northover et al. (2017a) excluded studies that reported aggregated data. Taken alongside our meta-analysis, and the research reviewed in Chapter 1 showing pronounced differences in the way people process information when we are or feel watched, Bradley et al.'s results suggest that it may be too soon to reach a definitive judgement on whether there is or is not a watching eyes effect. It may be an unsatisfying conclusion, but further research is clearly needed.

Referring back to our consideration of the 'mere presence' literature gives us further reason for suspending judgment on the effect of perceived surveillance on prosocial behaviour. The results of studies seeking to identify how human behaviour changes when we are watched directly, by another person, are eclectic and inconsistent. Given that the results of direct gaze are so inconsistent and context dependent, perhaps we should not have expected consistent responses to eye cues. For example, if direct gaze can both reduce performance (Forgas et al., 1980) and improve it (Laird et al., 2018), depending on the context, the nature of observer, the task and the nature of the person observed, then it is equally unlikely that eye cues will have consistent effects. We might all attend to gaze preferentially, and show consistent unconscious neural responses to direct gaze, but the field is not yet able to establish boundary conditions for direct gaze effects, let alone those for perceived surveillance.

That said, it might be that field experiments offer more consistent effects, although for different reasons in each case. Studies 5 & 6 did not demonstrate this with confidence. Nevertheless, they do both contribute to the literature, albeit in different ways. The Hereford field study (Study 5), for example, adds to the literature by suggesting that signs with eyes and text may reduce crime, perhaps with large effects comparable to the 62% reduction found in Nettle et al. (2012). The Ealing field study (Study 6) contributes via the careful design of the study - it is the first field experiment in the ‘watching eyes’ canon that has sought to differentiate the effects of different primes in crime prevention signage - and it demonstrates the utility of collaborating with local authorities. In both of these regards, Study 5 & 6 point the way for future ‘watching eyes’ researchers to test over wider areas on meaningful outcome variables, with more careful designs in ecologically valid environments.

Finally, while our hypothesis that eye cues might have more consistent effects on antisocial behaviour than prosocial was not borne out by our studies, the reduction in antisocial behaviour in the presence of eye cues found in our updated meta-analytic effect size remains significant and, at 29%, meaningful in public policy terms. From a policy perspective, low cost watching eyes signs interventions remain justifiable based on the research presented herein.

## 7.5. Limitations

The major limitations in the research presented are in the two field studies. The Hereford field study (Study 5), could not separate the effects of eye cues from the effect of the text or logos in the signs, and lacked a baseline to control for wider fluctuations in the crime rate. The Ealing field study (Study 6) was confounded both by the poor choice of intervention locations and by the fact that there was no cycle crime at all across the Borough in the 6 months after the signs were erected. Thus, the results in both field studies cannot provide definitive evidence for or against the effect of eye cues on antisocial behaviour.

The major potential limitation in our research in both the laboratory and field is the possibility that participants simply didn’t feel watched by the eye cues or the presence of cameras. This would, of course, explain the failure to find an effect of surveillance cues on public self-awareness.

Inversely, in the laboratory and online studies, it might be that participants felt just as 'watched' in the control condition as when surveillance cues were present. In Chapter 4 we suggested that it may not actually be possible to prevent participants feeling watched when taking part in a psychology experiment under laboratory conditions. Noah, Schul, and Mayo (2018) have tested this hypothesis and found some preliminary evidence to this end, while Bolton & Zwick (1995) eschewed the use of computers in laboratory experiments because they judged that participants would know that the computer might record everything they did and so confound any results. This concern might be overstated – Bradley et al.'s review of the effects of observability on prosociality suggests that laboratory experiments find more consistent effects than field studies. Regardless, both the gap between the results of our Hereford field experiment and those undertaken in the laboratory and the wider import of the question for much – perhaps all - laboratory-based psychological science, suggest this is an avenue worthy of further investigation.

If the language of the eyes and complexity of the messages the eyes can send is as complex as the evidence reviewed in Chapter 1 suggests, it is likely that the types of eyes selected matters. Studies 1, 2, 3, and 4, all used eye cues that were neutral, whereas our field studies used angry eyes. It is possible that the type of eye cue used has an effect, something the research presented in this thesis failed to address.

Due to the null effects found in the laboratory studies, it was impossible to determine whether the shape and colouration of the eye are particularly important in surveillance effects – as might be the case if hypersensitivity to surveillance began as a hypersensitivity to the eyes of predators, prey and other humans. The design of each study had therefore been intended to detect differences between the presence of eye cues compared to the presence of a camera, but the null results prevented a reliable comparison.

A larger criticism of the wider research on watching eyes can be levelled at its unhelpful heterogeneity. The plethora of different outcome variables and experimental designs creates a heterogeneity that makes meta-analysis and identifying boundary conditions difficult. For example, our own meta-analysis covered everything from littering to engine idling, to cycle theft and fly-tipping. Similarly, Northover et al.'s (2017a) meta-analysis is just as heterogenous. Even if we survey just some of the dictator game designs included in their study, we can see just how heterogenous watching eyes studies have been: one was conducted in darkness (Tane & Takezawa, 2011), the form of eye cues varied across studies from stylized eyes used by some (Haley & Fessler, 2005; Cai et al. 2015; Vogt et al. 2015), to painted eyes (Oda et al 2011), a 3-dot pattern (Rigdon, Ishii, Watabe, & Kitayama, 2009), cartoon eyes (Keller & Pfattheicher, 2011; Pfattheicher & Keller, 2015), neutral eyes (Nettle et al., 2013; Matsugasaki et al. 2015) even a design with 50 different types of eyes (Raihani & Bshary, 2012). One even included a statue of Erasmus' head as an 'eye cue' (Baillon et al., 2013). Nor was it just the types of eyes that differed either. One study, for example, selected the recipient in the dictator game to ensure he or she was from an in-group or out-group in relation to the dictator (Mifune et al., 2010), and others varied the length of exposure (Sparks & Barclay, 2013; White, 2015). Treating studies that vary as much as those in analysed in Chapter 3, and those in Northover et al.'s analysis as if they were the same is a stretch and may account for the widely varying effect sizes.

## 7.6. Further Research

An obvious follow on from the discussion above is the need for a systematic review and meta-analysis of watching eyes effects in field experiments to compare against a similar analysis of effects in laboratory conditions. With hindsight, this would have been a useful starting point for the research presented in this thesis. Similarly, there is, to the author's knowledge, no systematic review or meta-analysis of the effect of the presence of cameras on behaviour. Following on from the exploratory analyses conducted in this thesis a synthesis of camera effects would be informative both for surveillance research and research in public policy.

Relatedly, brain imaging studies have been able to show that direct gaze and thinking about God have similar neural signatures, but as yet there has been no investigation as to whether responses to other forms of perceived surveillance are similar. For example, do people use different neural pathways to process information when cued to feel the presence of others through the use of social media? Are the neural pathways used for action control different when we are in the presence of a surveillance camera? Similarly, eye tracking studies could help to determine whether we pay greater attention to cameras or other surveillance stimuli in comparison to other objects, much as we do with eyes.

Similarly, drawing on our previous observation, a study of the possible effects of different types of eye cues might help to better inform the understanding of eye cues. Given the range of emotional messages that our eyes can communicate, as described in Chapter 1, it might be that certain types of eyes have more effect on particular behaviours than others. For example, the angry eyes used in the field studies 5 & 6 might be more effective in inducing behavioural change than the neutral eyes used in studies 1-4.

As noted above, a more systematic programme of field research would be beneficial for understanding the effect of eye cues on behaviour. Collaborations such as those that enabled studies 5 & 6 are beneficial for both parties and enable testing in a more ecologically valid environment, away from the possible confounding variables of the laboratory such as the possible difficulty of preventing participants feeling watched or the generally low stakes for participants in the decisions they make in economic games.

Finally, one approach to counteract the problem of heterogeneity in this field of research is to take a team science approach and build cross-laboratory collaborations with other researchers to more systematically test and analyse watching eyes effects with large sample sizes. This is particularly important in relation to the current "reproducibility crisis" in psychological science and would help to build a more reliable body of evidence on the impact of surveillance on prosocial and anti-social behaviour.

## 7.7. Conclusion

This thesis examines the effect of surveillance on behaviour across 6 empirical studies exploring the effect of eye cues or the presence of a camera on behaviour. It finds no effect of eye cues or the presence of cameras on either prosocial behaviour in the first laboratory study, nor on antisocial behaviour in the second laboratory study nor in two further online studies. In the first field experiment there were consistent effects of signs which incorporated eye cues on antisocial behaviour across nine different locations and outcome measures, three of which were statistically significant and equated to a 27% reduction in shoplifting, a 100% reduction in cycle crime and a 100% reduction in fly-tipping. A large scale and carefully designed replication and extension of Nettle et al.'s study on the effect of watching eyes on cycle crime in Ealing was conducted but was unable to produce reliable results due to unexpected fluctuations in the crime rate Borough wide. A meta-analysis of all the available studies on the effect of eye cues on antisocial behaviour, including Studies 3, 4, 5, and 6 reported herein, found a 29% reduction in the risk of antisocial behaviour in the presence of eyes. Deriving our figures from the UK's costs of crime survey we noted that a 1% reduction in crime might be said to equate to a saving of at least £350,000,000 (Brand & Price, 2000) so while our results are inconclusive, they are far from trivial. In conclusion, collaborative partnerships with carefully evaluative designs conducted in the field may be the best way to test the effectiveness of watching eyes and inform wider surveillance research whilst, potentially, offering considerable savings to research partners in public policy with responsibility for crime reduction.

Returning to the bigger questions raised in Chapter 1 we can see from our evidence syntheses in the latter half of that chapter that Sartre may well have been right: *"We cannot perceive the world and at the same time apprehend a look fastened upon us; it must be either one or the other"* (Sartre, 1978, p. 258). "*The look*" in Sartre's words – or gaze in ours - changes the neural pathways for motor control and object perception, and facilitates perspective taking. In these respects, Chapter 1's evidence synthesis partially upholds the ideas of Foucault too, suggesting that being watched does have an effect on mental processing. However, while there is evidence, both in Study 5 in this thesis and the wider research field on eye cue effects, that feeling watched may change behaviour the evidence is not yet sufficient to uphold Bentham's suggestion that this alone is enough to make us more rule-abiding, nor Foucault's idea that it would lead to self-censorship. It seems likely that any effect is just as context dependent and variable as the effects of direct gaze on behaviour.

# Bibliography

Aarts, A. A., Anderson, J. E., Anderson, C. J., Attridge, P. R., Attwood, A., Axt, J., . . . Collaboration, O. S. (2015). Estimating the reproducibility of psychological science. *Science, 349*(6251), aac4716. doi: 10.1126/science.aac4716

Abbink, K., & Herrmann, B. (2011). The Moral Costs of Nastiness. *Economic Inquiry, 49*(2), 631-633. doi:10.1111/j.1465-7295.2010.00309.x

Abbink, K., & Sadrieh, A. (2009). The pleasure of being nasty. *Economics Letters, 105*(3), 306-308. doi:10.1016/j.econlet.2009.08.024

Abrams, D., & Manstead, A. S. (1981). A test of theories of social facilitation using a musical task. *British Journal of Social Psychology, 20*(4), 271-278.

Abrams, D., Wetherell, M., Cochrane, S., Hogg, M. A., & Turner, J. C. (1990). Knowing what to think by knowing who you are: Self-categorization and the nature of norm formation, conformity and group polarization*. *British Journal of Social Psychology, 29*(2), 97-119. doi:10.1111/j.2044-8309.1990.tb00892.x

Abrams, D. (1999). Social identity, social cognition, and the self: The flexibility and stability of self-categorization. In *Social identity and social cognition.* (pp. 197-229). Malden: Blackwell Publishing.

Adams, R. B., Jr., & Kleck, R. E. (2003). Perceived gaze direction and the processing of facial displays of emotion. *Psychological Science, 14*(6), 644-647. doi:10.1046/j.0956-7976.2003.psci_1479.x

Adams, R. B., Jr., & Kleck, R. E. (2005). Effects of direct and averted gaze on the perception of facially communicated emotion. *Emotion, 5*(1), 3-11. doi:10.1037/1528-3542.5.1.3

Aimone, J. A., & Houser, D. (2012). What you don't know won't hurt you: a laboratory analysis of betrayal aversion. *Experimental Economics, 15*(4), 571-588. doi:10.1007/s10683-012-9314-z

Alexandrie, G. (2017). Surveillance cameras and crime: a review of randomized and natural experiments. *Journal of Scandinavian Studies in Criminology and Crime Prevention, 18*(2), 210-222. doi:10.1080/14043858.2017.1387410

Altman, D. G., & Bland, J. M. (2011). How to obtain the P value from a confidence interval. *BMJ, 343*. doi:10.1136/bmj.d2304

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*: American Psychiatric Pub.

Aravindan, A., & Geddie, J. (2018). Singapore to test facial recognition on lampposts, stoking privacy fears. *Reuters*. Retrieved from https://www.reuters.com/article/us-singapore-surveillance/singapore-to-test-facial-recognition-on-lampposts-stoking-privacy-fears-idUSKBN1HK0RV

Archer, A. K. (2017). *"I Made a Choice": Exploring the Persuasion Tactics Used by Online Romance Scammers in Light of Cialdini's Compliance Principles.* (MSc Criminology), Regis University, Online. Retrieved from https://epublications.regis.edu/theses/823

Ariel, B., Farrar, W. A., & Sutherland, A. (2015). The Effect of Police Body-Worn Cameras on Use of Force and Citizens' Complaints Against the Police: A Randomized Controlled Trial. *Journal of Quantitative Criminology, 31*(3), 509-535. doi:10.1007/s10940-014-9236-3

Ariel, B., Sutherland, A., Henstock, D., Young, J., Drover, P., Sykes, J., . . . Henderson, R. (2016). Wearing body cameras increases assaults against officers and does not reduce police use of force: Results from a global multi-site experiment. *European Journal of Criminology, 13*(6), 744-755. doi:10.1177/1477370816643734

Ariel, B., Sutherland, A., Henstock, D., Young, J., Drover, P., Sykes, J., . . . Henderson, R. (2017). "Contagious Accountability": A Global Multisite Randomized Controlled Trial on the Effect of Police Body-Worn Cameras on Citizens' Complaints Against the Police. *Criminal Justice and Behavior, 44*(2), 293-316. doi:10.1177/0093854816668218

Armstrong, K. (2004). *A short history of myth* (Vol. 1). Canongate Books.

Arrington, M. (12 January 2010). OK You Luddites, Time to Chill Out On Facebook Over Privacy. Retrieved from https://techcrunch.com/2010/01/12/ok-you-luddites-time-to-chill-on-facebook-over-privacy/

Ashton, M. C., & Lee, K. (2009). The HEXACO-60: a short measure of the major dimensions of personality. *J Pers Assess, 91*(4), 340-345. doi:10.1080/00223890902935878

Ashton, M. C., & Lee, K. (2010). Trait and source factors in HEXACO-PI-R self- and observer reports. *European Journal of Personality, 24*(3), 278-289. doi:10.1002/per.759

Ashton, M. C., & Lee, K. (2018). How Well Do Big Five Measures Capture HEXACO Scale Variance? *Journal of Personality Assessment*, 1-7. doi:10.1080/00223891.2018.1448986

Association, A. P. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*: American Psychiatric Pub.

Axelrod, R., & Hamilton, W. D. (1981). The Evolution of Cooperation. *Science, 211*(4489), 1390-1396.

Axelrod, V., Bar, M., & Rees, G. (2015). Exploring the unconscious using faces. *Trends Cogn Sci, 19*(1), 35-45. doi:10.1016/j.tics.2014.11.003

Bağlı, H., & Öğüt, Ş. T. (2009). Towards an Analysis of the Signs of the 'Unknown': Objects with Rituals in Turkish Culture. *The Design Journal, 12*(3), 365-382. doi:10.2752/146069209X12530928086441

Baillon, A., Selim, A. I., & van Dolder, D. (2013). On the social nature of eyes: The effect of social cues in interaction and individual choice tasks. *Evolution and Human Behavior, 34*(2), 146-154. doi:10.1016/j.evolhumbehav.2012.12.001

Balmforth, T. (24 November 2012). Cops, Cars, And Videotape: Russians Embrace Dash-Cam Craze. *Radio Free Europe / Radio Liberty*. Retrieved from https://www.rferl.org/a/dash-cams-russia-fighting-corruption-and-scams-car-crashes/24780355.html

Bar-Yosef, O. (1998). On the Nature of Transitions: the Middle to Upper Palaeolithic and the Neolithic Revolution. *Cambridge Archaeological Journal, 8*(2), 141-163. doi:10.1017/S0959774300001815

Barclay, P. (2013). Strategies for cooperation in biological markets, especially for humans. *Evolution and Human Behavior, 34*(3), 164-175. doi:10.1016/j.evolhumbehav.2013.02.002

Barker, G. (2006). *The agricultural revolution in prehistory : why did foragers become farmers?* Oxford: Oxford University Press.

Barlow, G. W. (1972). The attitude of fish eye-lines in relation to body shape and to stripes and bars. *Copeia*, 4-12.

Baron-Cohen, S. (1997a). How to build a baby that can read minds: Cognitive mechanisms in mindreading. In S. Baron-Cohen (Ed.), *The maladapted mind: Classic readings in evolutionary psychopathology* (pp. 207-239). Hove: Psychology Press.

Baron-Cohen, S. (1997b). *Mindblindness: An essay on autism and theory of mind*: MIT press.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Test Revised Version: A Study with Normal Adults, and Adults with Asperger Syndrome or High-functioning Autism. *Journal of Child Psychology and Psychiatry, 42*(2), 241-251. doi:10.1111/1469-7610.00715

Basildon, Canvey, & Southend Echo. (2013). Poster trial on c2c line to target bike thieves. *Basildon, Canvey, Southend Echo*. Retrieved from http://www.echo-news.co.uk/news/local_news/10373770.Poster_trial_on_c2c_line_to_target_bike_thieves/?ref=nt

Bateson, M., Callow, L., Holmes, J. R., Roche, M. L. R., & Nettle, D. (2013). Do Images of 'Watching Eyes' Induce Behaviour That Is More Pro-Social or More Normative? A Field Experiment on Littering. *PLoS ONE, 8*(12), e82055. doi:10.1371/journal.pone.0082055

Bateson, M., Nettle, D., & Roberts, G. (2006). Cues of being watched enhance cooperation in a real-world setting. *Biology letters, 2*(3), 412-414. doi:10.1098/rsbl.2006.0509

Bateson, M., Robinson, R., Abayomi-Cole, T., Greenlees, J., O'Connor, A., & Nettle, D. (2015). Watching eyes on potential litter can reduce littering: evidence from two field experiments. *PeerJ, 3*, e1443. doi:10.7717/peerj.1443

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of general psychology, 5*(4), 323.

Baumeister, R. F., Zhang, L., & Vohs, K. D. (2004). Gossip as cultural learning. *Review of general psychology, 8*(2), 111.

Bayliss, A. P., Paul, M. A., Cannon, P. R., & Tipper, S. P. (2006). Gaze cuing and affective judgments of objects: I like what you look at. *Psychonomic bulletin & review, 13*(6), 1061-1066. doi:10.3758/Bf03213926

BBC News. (2013). Staring eyes 'deter' Newcastle University bike thieves. Retrieved from http://www.bbc.co.uk/news/uk-england-tyne-22270052

BBC News. (2015). Tax crackdown by HMRC unit nets 60% more money, report says. Retrieved from BBC News website: http://www.bbc.co.uk/news/business-30973191

Becchio, C., Bertone, C., & Castiello, U. (2008). How the gaze of others influences object processing. *Trends in Cognitive Sciences, 12*(7), 254-258. doi:10.1016/j.tics.2008.04.005

Bednar, J. A., & Miikkulainen, R. (2003). Learning innate face preferences. *Neural Compututation, 15*(7), 1525-1557. doi:10.1162/089976603321891792

Bentham, J. (1791). *Panopticon or the inspection house* (Vol. 2).

Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological Studies of Face Perception in Humans. *Journal of Cognitive Neuroscience, 8*(6), 551-565. doi:10.1162/jocn.1996.8.6.551

Bering, J. M. (2006). The cognitive science of souls: Clarifications and extensions of the evolutionary model. *Behavioral and brain sciences, 29*(5), 486-493.

Bering, J. M., McLeod, K., & Shackelford, T. K. (2005). Reasoning about dead agents reveals possible adaptive trends. *Human Nature, 16*(4), 360-381. doi:10.1007/s12110-005-1015-2

Bering, J. M., & Shackelford, T. K. (2004). The Causal Role of Consciousness: A Conceptual Addendum to Human Evolutionary Psychology. *Review of general psychology, 8*(4), 227-248. doi:10.1037/1089-2680.8.4.227

Beyfus, T. A., Dawson, N. L., Danner, C. H., Rawal, B., Gruber, P. E., & Petrou, S. P. (2016). The use of passive visual stimuli to enhance compliance with handwashing in a perioperative setting. *American journal of infection control, 44*(5), 496-499. doi:10.1016/j.ajic.2015.12.021

Big Brother Watch (Producer). (2012). The Price of Privacy: How local authorities spent £515m on CCTV in four years. Retrieved from https://www.bigbrotherwatch.org.uk/files/priceofprivacy/Price_of_privacy_2012.pdf

Big Brother Watch (Producer). (2016). Are they still watching? The cost of CCTV. Retrieved from https://www.bigbrotherwatch.org.uk/wp-content/uploads/2016/02/Are-They-Still-Watching.pdf

Black, D. (2014). *Toward a general theory of social control: Fundamentals* (Vol. 1): Academic Press.

Blass, T. (1991). Understanding Behavior in the Milgram Obedience Experiment - the Role of Personality, Situations, and Their Interactions. *Journal of Personality and Social Psychology, 60*(3), 398-413. doi: 10.1037/0022-3514.60.3.398

Blass, T. (2012). A Cross-Cultural Comparison of Studies of Obedience Using the Milgram Paradigm: A Review. In (Vol. 6, pp. 196-205). Oxford, UK.

Blest, A. D. (1957). The Function of Eyespot Patterns in the Lepidoptera. *Behaviour, 11*(2-3), 209-258. doi:10.1163/156853956X00048

Blum, B. (7 June 2018). The Lifespan of a Lie: The most famous psychology study of all time was a sham. Why can’t we escape the Stanford Prison Experiment? Retrieved from https://medium.com/s/trustissues/the-lifespan-of-a-lie-d869212b1f62

Bohnet, I., Greig, F., Herrmann, B., & Zeckhauser, R. (2008). Betrayal Aversion: Evidence from Brazil, China, Oman, Switzerland, Turkey, and the United States. *American Economic Review, 98*(1), 294-310. doi: 10.1257/aer.98.1.294

Bolton, G. E., & Zwick, R. (1995). Anonymity versus Punishment in Ultimatum Bargaining. *Games and Economic Behavior, 10*(1), 95-121. doi:https://doi.org/10.1006/game.1995.1026

Bourrat, P., Baumard, N., & McKay, R. (2011). Surveillance cues enhance moral condemnation. *Evol Psychol, 9*(2), 193-199. doi:10.1177/147470491100900206

Boyd, R., & Richerson, P. J. (1988). The evolution of reciprocity in sizable groups. *Journal of theoretical Biology, 132*(3), 337-356.

Boyer, P. (2003). Religious thought and behaviour as by-products of brain function. *Trends in Cognitive Sciences, 7*(3), 119-124. doi:https://doi.org/10.1016/S1364-6613(03)00031-7

Bradley, A., Lawrence, C., & Ferguson, E. (2018). Does observability affect prosociality? *Proceedings of the Royal Society B: Biological Sciences, 285*(1875). doi:10.1098/rspb.2018.0116

Brañas-Garza, P., Espín, A., Exadaktylos, F., & Herrmann, B. (2014). Fair and unfair punishers coexist in the Ultimatum Game. *Scientific Reports (Nature Publisher Group), 4*, 6025. doi:10.1038/srep06025

Brand, S., & Price, R. (2000). *The Economic and Social Costs of Crime* (217). Retrieved from Online: http://webarchive.nationalarchives.gov.uk/20110218135832/rds.homeoffice.gov.uk/rds/pdfs/hors217.pdf

Brenner, B. (2017). Know the risks of Amazon Alexa and Google Home. *Naked Security*. Retrieved from https://nakedsecurity.sophos.com/2017/01/27/data-privacy-day-know-the-risks-of-amazon-alexa-and-google-home/

Brien, T. J., & Dunlap, W. P. (1975). Tonic immobility in the blue crab ( Callinectes sapidus, Rathbun): It's relation to threat of predation. *Journal of Comparative and Physiological Psychology, 89*(1), 86-94. doi:10.1037/h0076425

Brudermann, T., Bartel, G., Fenzl, T., & Seebauer, S. (2015). Eyes on social norms: A field study on an honor system for newspaper sale. *Theory and Decision, 79*(2), 285-306. doi:10.1007/s11238-014-9460-1

Brugger, P. (2001). From haunted brain to haunted science: A cognitive neuroscience view of paranormal and pseudoscientific thought. *Hauntings and poltergeists: Multidisciplinary perspectives, ed. J. Houran & R. Lange*, 195-213.

Buchanan, J., McMahon, M., Simpson, M., & Wilson, B. (2017). Smile, Dictator, You're on Camera. *Southern Economic Journal, 84*(1), 52-65. doi:10.1002/soej.12214

Burger, J. M. (2009). Replicating Milgram: Would people still obey today? *Am Psychol, 64*(1), 1-11. doi:10.1037/a0010932

Burgess, W. E. (1978). Butterflyfishes of the world; a monograph of the family Chaetodontidae.

Burghardt, G. M., & Greene, H. W. (1988). Predator Simulation and Duration of Death Feigning in Neonate Hognose Snakes. *Animal Behaviour, 36*(6), 1842-1844. doi: 10.1016/S0003-3472(88)80127-1

Burnham, T. C. (2003). Engineering altruism: a theoretical and experimental investigation of anonymity and gift giving. *Journal of economic behavior & organization, 50*(1), 133-144. doi:10.1016/S0167-2681(02)00044-6

Burnham, T. C., & Hare, B. (2007). Engineering Human Cooperation : Does Involuntary Neural Activation Increase Public Goods Contributions? *Hum Nat, 18*(2), 88-108. doi:10.1007/s12110-007-9012-2

Burra, N., Hervais-Adelman, A., Kerzel, D., Tamietto, M., de Gelder, B., & Pegna, A. J. (2013). Amygdala Activation for Eye Contact Despite Complete Cortical Blindness. *The Journal of Neuroscience, 33*(25), 10483-10489. doi:10.1523/jneurosci.3994-12.2013

Bush, S. S., Erlich, A., Prather, L., & Zeira, Y. (2016). The Effects of Authoritarian Iconography: An Experimental Test. *Comparative Political Studies, 49*(13), 1704-1738. doi:10.1177/0010414016633228

Bushman, B. J. (1984). Perceived symbols of authority and their influence on compliance. *Journal of Applied Social Psychology, 14*(6), 501-508.

Buss, A. H. (1980). *Self-consciousness and social anxiety*: Freeman.

Cacioppo, J. T., & Petty, R. E. (1986). Social processes. In M. G. H. Coles, E. Donchin, & S. Porges (Eds.), *Psychophysiology: Systems, Processes, and Applications* (pp. 646–679). New York: Guilford Press.

Cai, W., Huang, X., Wu, S., & Kou, Y. (2015). Dishonest behavior is not affected by an image of watching eyes. *Evolution and Human Behavior, 36*(2), 110-116. doi:http://dx.doi.org/10.1016/j.evolhumbehav.2014.09.007

Carabas, M. (2018). *Horus: The History and Legacy of the Ancient Egyptian God Who Was the Son of Isis and Osiris*: Create Space Independent Publishing Platform.

Cayford, M., & Pieters, W. (2018). The effectiveness of surveillance technology: What intelligence officials are saying. *The Information Society, 34*(2), 88-103. doi:10.1080/01972243.2017.1414721

Chan, T. F. (1 May 2018). China is monitoring employees' brain waves and emotions — and the technology boosted one company's profits by $315 million. Retrieved from http://uk.businessinsider.com/china-emotional-surveillance-technology-2018-4

Chavez, N. (2 December 2017). Arkansas judge drops murder charge in Amazon Echo case. *CNN International Edition +*. Retrieved from https://edition.cnn.com/2017/11/30/us/amazon-echo-arkansas-murder-case-dismissed/index.html

Chen, S. (17 July 2018). 'Forget the Facebook leak': China is mining data directly from workers' brains on an industrial scale. *South China Morning Post*. Retrieved from https://www.scmp.com/news/china/society/article/2143899/forget-facebook-leak-china-mining-data-directly-workers-brains

Chowdhury, S., Jeon, J. Y., & Saha, B. (2014). *Eye-image in Experiments: Social Cue or Experimenter Demand Effect?* Retrieved from Norwich, UK.: http://EconPapers.repec.org/RePEc:uea:aepppr:2012_67

Cialdini, R. B. (2007). *Influence: The psychology of persuasion*: Collins New York.

Cluley, G. (21 March 2015). Hey Siri, How Come Strangers Are Listening To My Private Messages? Retrieved from https://www.tripwire.com/state-of-security/security-awareness/siri-privacy/

Coe, R. (2002). *It's the effect size, stupid: What effect size is and why it is important*. Paper presented at the Annual Conference of the British Educational Research Association, University of Exeter, UK. http://www.leeds.ac.uk/educol/documents/00002182.htm

Cohen, J. (1992). A power primer. *Psychological bulletin, 112*(1), 155.

Conrad, K. (1958). Die beginnende Schizophrenie: Versuch einer gestaltanalyse des Wahns [The onset of schizophrenia: An experimental analysis of creative madness]. Thieme: Stuttgart, Germany.

Conty, L., Dezecache, G., Hugueville, L., & Grezes, J. (2012). Early binding of gaze, gesture, and emotion: neural time course and correlates. *J Neurosci, 32*(13), 4531-4539. doi:10.1523/JNEUROSCI.5636-11.2012

Conty, L., George, N., & Hietanen, J. K. (2016). Watching Eyes effects: When others meet the self. *Consciousness and cognition, 45*, 184-197. doi:10.1016/j.concog.2016.08.016

Conty, L., N'Diaye, K., Tijus, C., & George, N. (2007). When eye creates the contact! ERP evidence for early dissociation between direct and averted gaze motion processing. *Neuropsychologia, 45*(13), 3024-3037. doi:10.1016/j.neuropsychologia.2007.05.017

Conty, L., Russo, M., Loehr, V., Hugueville, L., Barbu, S., Huguet, P., . . . George, N. (2010). The mere perception of eye contact increases arousal during a word-spelling task. *Soc Neurosci, 5*(2), 171-186. doi:10.1080/17470910903227507

Coss, R. G. (1978). Perceptual Determinants of Gaze Aversion By the Lesser Mouse Lemur ( Microcebus Murinus), the Role of Two Facing Eyes. *Behaviour, 64*(F0020003), 248-269.

Coss, R. G. (1979). Delayed plasticity of an instinct: recognition and avoidance of 2 facing eyes by the jewel fish. *Dev Psychobiol, 12*(4), 335-345. doi:10.1002/dev.420120408

Cosmides, L., & Tooby, J. (1997, 13 January 1997). Evolutionary Psychology: A Primer. Retrieved from http://www.cep.ucsb.edu/primer.html

Cott, H. B. (1940). *Adaptive coloration in animals*. London.

Cott, H. B. (1957). *Adaptive coloration in animals*. London: Methuen.

Cottrell, N. B. (1968). Performance in the presence of other human beings: mere presence, audience and affiliation effects. In E. C. Samuel, R. A. Hoppe, & G. A. Milton (Eds.), *Social facilitation and imitative behaviour*. Boston: Allyn & Bacon.

Couts, A. (30 October 2012). State of the Web: Who Killed Privacy? You Did. Retrieved from https://www.digitaltrends.com/opinion/state-of-the-web-who-killed-privacy/

Cox, J., & Li, D. (2012). Do I care if you know I betrayed you? *Economics Bulletin, 32*(4), 2839-2848.

Crawley, M. J. (2012). *The R Book* (2nd ed.). Somerset: John Wiley & Sons.

Curry, A. (2008, November 2008). Gobekli Tepe: The World's First Temple? *Smithsonian Magazine*. Retrieved from http://www.smithsonianmag.com/history/gobekli-tepe-the-worlds-first-temple-83613665/

Darvill, T., Marshall, P., Pearson, M. P., & Wainwright, G. (2012). Stonehenge remodelled. *Antiquity, 86*(334), 1021-1040.

De Bona, S., Valkonen, J. K., López-Sepulcre, A., & Mappes, J. (2015). Predator mimicry, not conspicuousness, explains the efficacy of butterfly eyespots. *Proceedings of the Royal Society B: Biological Sciences, 282*(1806). doi:10.1098/rspb.2015.0202

Dear, K. P., & Ellen, M. (2018). Joy of Destruction: A Research Tool for Exploring Anti-Social Behaviour. Online: Open Science Framework. Retrieved from osf.io/f2vsx

Deci, E. L., & Ryan, R. M. (2002). *Handbook of self-determination research*: University Rochester Press.

Delton, A. W., Krasnow, M. M., Cosmides, L., & Tooby, J. (2011). Evolution of direct reciprocity under uncertainty can explain human generosity in one-shot encounters. *Proceedings of the national Academy of Sciences, 108*(32), 13335-13340. doi:10.1073/pnas.1102131108

Demolliens, M., Isbaine, F., Takerkart, S., Huguet, P., & Boussaoud, D. (2017). Social and asocial prefrontal cortex neurons: a new look at social facilitation and the social brain. *Social cognitive and affective neuroscience, 12*(8), 1241-1248. doi:10.1093/scan/nsx053

de Winter, J., & Happee, R. (2013). Why Selective Publication of Statistically Significant Results Can Be Effective. *PLoS ONE, 8*(6), e66463. doi:10.1371/journal.pone.0066463

Diehl, P. F. (1993). International peacekeeping (Perspectives on security). *Baltimore: Johns Hopkins Univ. Press, 3*, 34.

Dorn, A. W. (2014). *Air power in UN operations: wings for peace*: Ashgate Publishing, Ltd.

Doyle, M. W., & Sambanis, N. (2000). International peacebuilding: A theoretical and quantitative analysis. *American political science review, 94*(4), 779-801.

Draper, R. (February 2018). They are watching you - and everything else on the planet. *National Geographic*.

Dreber, A., Ellingsen, T., Johannesson, M., & Rand, D. G. (2013). Do people care about social context? Framing effects in dictator games. *Experimental Economics, 16*(3), 349-371. doi:10.1007/s10683-012-9341-9

Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., & Baron-Cohen, S. (1999). Gaze perception triggers reflexive visuospatial orienting. *Visual cognition, 6*(5), 509-540. doi:10.1080/135062899394920

Dunbar, R. I. M. (1992). Neocortex size as a constraint on group size in primates. *Journal of Human Evolution, 22*(6), 469-493. doi:https://doi.org/10.1016/0047-2484(92)90081-J

Dunbar, R. I. M. (1993). Coevolution of neocortical size, group size and language in humans. *Behavioral and brain sciences, 16*(4), 681-694. doi:10.1017/S0140525X00032325

Dunbar, R. I. M. (1996). Determinants of Group Size in Primates: A General Model. *Proceedings of the British Academy, 88*, 33-51.

Dunbar, R. I. M. (1998). The social brain hypothesis. *Evolutionary anthropology: Issues, news, and reviews, 6*(5), 178-190. doi: 10.1002/(SICI)1520-6505(1998)6:5<178::AID-EVAN5>3.0.CO;2-8

Dunbar, R. I. M. (2004). Gossip in evolutionary perspective. *Review of general psychology, 8*(2), 100.

Dunbar, R. I. M. (2013). The origin of religion as a small scale phenomenon. *Religion, intolerance, and conflict: A scientific and conceptual investigation*, 48-66.

Dunbar, R. I. M. (2014). *Human evolution: A Pelican introduction*: Penguin UK.

Dunbar, R. I. M. (2017). What's missing from the scientific study of religion? *Religion, Brain & Behavior, 7*(4), 349-353. doi:10.1080/2153599X.2016.1249927

Dunbar, R. I. M., & Sosis, R. (2018). Optimising human community sizes. *Evolution and Human Behavior, 39*(1), 106-111. doi:https://doi.org/10.1016/j.evolhumbehav.2017.11.001

Dundes, A. (1981). *The evil eye: A casebook* (Vol. 2): Univ of Wisconsin Press.

Dupret, V., Sanchez, S., Goujet, D., Tafforeau, P., & Ahlberg, P. E. (2014). A primitive placoderm sheds light on the origin of the jawed vertebrate face. *Nature, 507*, 500. doi:10.1038/nature12980 https://www.nature.com/articles/nature12980#supplementary-information

Duval, S., & Tweedie, R. (2000). Trim and Fill: A Simple Funnel- Plot– Based Method of Testing and Adjusting for Publication Bias in Meta- Analysis. *Biometrics, 56*(2), 455-463. doi:10.1111/j.0006-341X.2000.00455.x

Economist Intelligence Unit. (2018). *Democracy Index 2017*. Retrieved from Online: https://www.eiu.com/topic/democracy-index

Efferson, C., Lalive, R., Richerson, P. J., McElreath, R., & Lubell, M. (2008). Conformists and mavericks: the empirics of frequency-dependent cultural transmission. *Evolution and Human Behavior, 29*(1), 56-64.

Eisenberger, N. I., Lieberman, M. D., & Williams, K. D. (2003). Does rejection hurt? An fMRI study of social exclusion. *Science, 302*(5643), 290-292.

Ellis, D., Harper, D., & Tucker, I. (2016). Experiencing the 'surveillance society'. *Psychologist Magazine, 29*(9), 682-685.

Emery, N. J. (2000). The eyes have it: the neuroethology, function and evolution of social gaze. *Neuroscience and Biobehavioral Reviews, 24*(6), 581-604. doi: 10.1016/S0149-7634(00)00025-7

Ernest-Jones, M., Nettle, D., & Bateson, M. (2011). Effects of eye images on everyday cooperative behavior: a field experiment. *Evolution and Human Behavior, 32*(3), 172-178. doi:10.1016/j.evolhumbehav.2010.10.006

Extra Services, K. B. T. (2015). Litterers: We're Watching You [Press release]. Retrieved from http://www.extraservices.co.uk/article/extra-msa-keep-britain-tidy/

Fathi, M., Bateson, M., & Nettle, D. (2014). Effects of watching eyes and norm cues on charitable giving in a surreptitious behavioral experiment. *Evolutionary Psychology, 12*(5), 878-887.

Faul, F., Erdfelder, E., Lang, A.G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175-191. doi:10.3758/BF03193146

FDR Group, & PEN American Center. (2013). *Chilling effects: NSA surveillance drives US writers to self-censor*. Retrieved from https://pen.org/chilling-effects/

Fehr, E., & Fischbacher, U. (2003). The nature of human altruism. *Nature, 425*(6960), 785-791. doi:10.1038/nature02043

Fehr, E., & Gachter, S. (2002). Altruistic punishment in humans. *Nature, 415*(6868), 137-140. doi:10.1038/415137a

Fehr, E., & Schneider, F. (2010). Eyes are on us, but nobody cares: are eye cues relevant for strong reciprocity? *Proceedings of the Royal Society B, 277*(1686), 1315-1323. doi:10.1098/rspb.2009.1900

Feinberg, M., Willer, R., & Schultz, M. (2014). Gossip and ostracism promote cooperation in groups. *Psychological science, 25*(3), 656-664.

Feng, Q., & Zhang, X. (2014). Eye gaze triggers reflexive attention shifts: Evidence from lateralised ERPs. *Brain Research, 1589*, 37-44. doi:https://doi.org/10.1016/j.brainres.2014.09.029

Fenigstein, A., Scheier, M. F., & Buss, A. H. (1975). Public and Private Self-Consciousness - Assessment and Theory. *Journal of consulting and clinical psychology, 43*(4), 522-527. doi: 10.1037/h0076760

Fenigstein, A., & Abrams, D. (1993). Self-Attention and the Egocentric Assumption of Shared Perspectives. *Journal of Experimental Social Psychology, 29*(4), 287-303. doi:https://doi.org/10.1006/jesp.1993.1013

Fenton, N., & Barassi, V. (2011). Alternative Media and Social Networking Sites: The Politics of Individuation and Political Participation. *The Communication Review, 14*(3), 179-196. doi:10.1080/10714421.2011.597245

Field, A. P., Field, Z., & Miles, J. (2012). *Discovering statistics using R*. London: Sage.

Field, A. P., & Wilcox, R. R. (2017). Robust statistical methods: A primer for clinical psychology and experimental psychopathology researchers. *Behaviour Research and Therapy, 98*, 19-38. doi:https://doi.org/10.1016/j.brat.2017.05.013

Fischer, P., Krueger, J. I., Greitemeyer, T., Vogrincic, C., Kastenmüller, A., Frey, D., . . . Kainbacher, M. (2011). The bystander-effect: A meta-analytic review on bystander intervention in dangerous and non-dangerous emergencies. In (Vol. 137, pp. 517-537). US: American Psychological Association.

Fiske, S., & Taylor, S. K. (2017). *Social Cognition: From Brains to Culture* (Third Edition). London: SAGE Publications Ltd.

Flanagan, P. (2013, 2 June 2013). Crime... the answer's staring you in the face: The striking new police poster that has slashed thefts by up to 40 per cent. *Daily Mail*. Retrieved from http://www.dailymail.co.uk/news/article-2334577/Crime--answers-staring-face-The-striking-new-police-poster-slashed-thefts-40-cent.html

Forgas, J. P., Brennan, G., Howe, S., Kane, J. F., & Sweet, S. (1980). Audience Effects on Squash Players' Performance. *The Journal of social psychology, 111*(1), 41-47. doi:10.1080/00224545.1980.9924271

Fortna, V. P. (2008). *Does peacekeeping work? Shaping belligerents' choices after civil war*: Princeton University Press.

Foucault, M. (1977). *Discipline and Punish: the Birth of the Prison* (A. Sheridan, Translator. 1st American Edition ed.). New York: Pantheon Books.

Fox, E., & Damjanovic, L. (2006). The eyes are sufficient to produce a threat superiority effect. *Emotion, 6*(3), 534.

Freiwald, W. A., & Tsao, D. Y. (2010). Functional compartmentalization and viewpoint generalization within the macaque face-processing system. *Science, 330*(6005), 845-851. doi:10.1126/science.1194908

Friesen, C. K., & Kingstone, A. (2003). Abrupt onsets and gaze direction cues trigger independent reflexive attentional effects. *Cognition, 87*(1), B1-B10. doi:10.1016/S0010-0277(02)00181-6

Froming, W. J., Walker, G. R., & Lopyan, K. J. (1982). Public and private self-awareness: When personal attitudes conflict with societal expectations. *Journal of Experimental Social Psychology, 18*(5), 476-487. doi:https://doi.org/10.1016/0022-1031(82)90067-1

Fujii, T., Takagishi, H., Koizumi, M., & Okada, H. (2015). The effect of direct and indirect monitoring on generosity among preschoolers. *Scientific reports, 5*. doi:10.1038/srep09025

Gächter, S. (2017). Occasional errors can benefit coordination. *Nature, 545*, 297. doi:10.1038/545297a

Gaiani, A., Rose, E., & Roberts, G. (2014). *How and Why Do Watching Eyes Enhance Pro-Social Behaviour*. http://research.ncl.ac.uk/expeditionresearchscholarships/postergalleries/2014posters/Annika Gaiani.pdf

Galizzi, M. M., & Navarro-Martinez, D. (2018). On the External Validity of Social Preference Games: A Systematic Lab-Field Study. *Management Science*. doi:10.1287/mnsc.2017.2908

Gallagher, S. (2016). Police ask: “Alexa, did you witness a murder?”. *Ars Technica UK*. https://arstechnica.com/tech-policy/2016/12/police-ask-alexa-did-you-witness-a-murder/

Gallup, A., Church, A. M., Miller, H., Risko, E. F., & Kingstone, A. (2016). Social Presence Diminishes Contagious Yawning in the Laboratory. *Scientific reports, 6*, 25045. doi:10.1038/srep25045

Gallup, G. G., Nash, R. F., & Ellison, A. L. (1971). Tonic immobility as a reaction to predation: Artificial eyes as a fear stimulus for chickens. *Psychonomic Science, 23*(1), 79-80. doi:10.3758/bf03336016

Geen, R. G. (1991). Social motivation. *Annual review of psychology, 42*(1), 377-399.

Geertz, C. (1975). On the Nature of Anthropological Understanding: Not extraordinary empathy but readily observable symbolic forms enable the anthropologist to grasp the unarticulated concepts that inform the lives and cultures of other peoples. *American Scientist, 63*(1), 47-53.

George, N., & Conty, L. (2008). Facing the gaze of others. *Neurophysiologie Clinique/Clinical Neurophysiology, 38*(3), 197-207. doi:10.1016/j.neucli.2008.03.001

Germain, S., Douillet, A.-C., & Dumoulin, L. (2011). The legitimization of CCTV as a policy tool: Genesis and stabilization of a socio-technical device in three French cities. *The British Journal of Criminology, 52*(2), 294-308.

Gervais, W. M. (2013). Perceiving Minds and Gods: How Mind Perception Enables, Constrains, and Is Triggered by Belief in Gods. *Perspectives on Psychological Science, 8*(4), 380-394. doi:10.1177/1745691613489836

Gervais, W. M. (2016). Putting PET-PEESE to the test. *Will Gervais Blog*. Retrieved from http://willgervais.com/blog/2015/6/25/putting-pet-peese-to-the-test-1

Gilovich, T., Medvec, V. H., & Savitsky, K. (2000). The spotlight effect in social judgment: an egocentric bias in estimates of the salience of one's own actions and appearance. *Journal of personality and social psychology, 78*(2), 211. doi:10.1037/0022-3514.78.2.211

Givens, D. B. (1978). The nonverbal basis of attraction: flirtation, courtship, and seduction. *Psychiatry, 41*(4), 346-359. doi:10.1080/00332747.1978.11023994

Goldstein, P. (2017). Will Alexa, Cortana, Siri and Google Be Taking Over Your Office in the Near Future? *Biz Tech*. Retrieved from Biz Tech website: https://biztechmagazine.com/article/2017/03/will-alexa-cortana-siri-and-google-be-taking-over-your-office-near-future

Golja, A. (2013). *How Cues of Being Watched Affect Moral Judgements.* Paper presented at the MEi: CogSci Conference 2013, Budapest.

Google. (2017, 17 April 2017). Google Privacy & Terms. Retrieved from https://www.google.co.uk/intl/en/policies/privacy/

Gordon, G. (2016). Monitoring Conflict to Reduce Violence: Evidence from a Satellite Intervention in Darfur.

Govern, J. M., & Marsch, L. A. (2001). Development and Validation of the Situational Self-Awareness Scale. *Consciousness and cognition, 10*(3), 366-378. doi:https://doi.org/10.1006/ccog.2001.0506

Grant, N. (2017). What a Murder in Arkansas Means for Your Digital Privacy. *Bloombery Technology*. Retrieved from https://www.bloomberg.com/news/articles/2017-06-13/what-a-murder-in-arkansas-means-for-your-digital-privacy

Griffin, A. (1 June 2017). Google Voice Search Records and Keeps Conversations People Have Around Their Phones - But the Files Can be Deleted. *The Independent*. Retrieved from http://www.independent.co.uk/life-style/gadgets-and-tech/news/google-voice-search-records-stores-conversation-people-have-around-their-phones-but-files-can-be-a7059376.html

Grossmann, T. (2017). The Eyes as Windows Into Other Minds: An Integrative Perspective. *Perspectives on Psychological Science, 12*(1), 107-121. doi:10.1177/1745691616654457

Gürerk, Ö., Irlenbusch, B., & Rockenbach, B. (2006). The Competitive Advantage of Sanctioning Institutions. *Science, 312*(5770), 108-111. doi:10.1126/science.1123633

Haley, K. J., & Fessler, D. M. (2005). Nobody's watching? Subtle cues affect generosity in an anonymous economic game. *Evolution and Human Behavior, 26*(3), 245-256.

Hamedani, M. G., Markus, H. R., & Fu, A. S. (2013). In the land of the free, interdependent action undermines motivation. *Psychological science*, 0956797612452864.

Hamilton, W. D. (1970). Selfish and Spiteful Behaviour in an Evolutionary Model. *Nature, 228*, 1218. doi:10.1038/2281218a0

Harari, Y. N. (2015). *Sapiens : a brief history of humankind*. London: Vintage Books.

Hartendorp, M., Van der Stigchel, S., Hooge, I., Mostert, J., de Boer, T., & Postma, A. (2013). The relation between gaze behaviour and categorisation: does where we look determine what we see? *Journal of Vision, 13*(6). doi:10.1167/13.6.6

Harter, S. (1978). Effectance Motivation Reconsidered. Toward a Developmental Model. *Human Development, 21*(1), 34-64.

Hass, R. G. (1984). Perspective Taking And Self-Awareness - Drawing An E On Your Forehead. *Journal of personality and social psychology, 46*(4), 788-798. doi:10.1037//0022-3514.46.4.788

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2002). Human neural systems for face recognition and social communication. *Biological psychiatry, 51*(1), 59-67. doi:10.1016/S0006-3223(01)01330-0

Heider, F., & Simmel, M. (1944). An Experimental Study of Apparent Behavior. *The American journal of psychology, 57*(2), 243-259. doi:10.2307/1416950

Heine, S. J. (2001). Self as cultural product: An examination of East Asian and North American selves. *Journal of personality, 69*(6), 881-905. doi: 10.1111/1467-6494.696168

Hempel, L., & Töpfer, E. (2004). CCTV in Europe.

Henderson, T. (2017). Virtual assistants hear everything, so watch what you say. I'm not kidding. *Network World*. Retrieved from http://www.networkworld.com/article/3190176/security/virtual-assistants-hear-everything-so-watch-what-you-say-i-m-not-kidding.html

Hennig, C. (1979). The effects of physical environment, time in captivity, and defensive distance on tonic immobility, freezing, and flight behaviors in Anolis carolinensis. *Animal Learning & Behavior, 7(1), 106-110.* doi:10.3758/BF03209666

Hennig, C. W. (1977). Effects of simulated predation on tonic immobility in Anolis carolinensis: The role of eye contact. *Bulletin of the Psychonomic Society, 9*(4), 239-242. doi:10.3758/bf03336987

Henrich, J., & Gil-White, F. J. (2001). The evolution of prestige: freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evolution and Human Behavior, 22*(3), 165-196. doi:https://doi.org/10.1016/S1090-5138(00)00071-4

Hermstruwer, Y., & Dickert, S. (2017). Sharing is daring: An experiment on consent, chilling effects and a salient privacy nudge. *International Review of Law and Economics, 51*, 38-49. doi:10.1016/j.irle.2017.06.001

Herrmann, B., Thöni, C., & Gächter, S. (2008). Antisocial Punishment Across Societies. *Science, 319*(5868), 1362-1367. doi:10.1126/science.1153808

Herrmann, E., Call, J., Hernandez-Lloreda, M. V., Hare, B., & Tomasello, M. (2007). Humans have evolved specialized skills of social cognition: The cultural intelligence hypothesis. *Science, 317*(5843), 1360-1366. doi:10.1126/science.1146282

Hertz, U., and Bahrami, B. (2018). Intrinsic Value Of Social Influence Over Others. *PsyArXiv*. doi:http://doi.org/10.17605/OSF.IO/6JM7T

Hess, E. H. (1975). *The tell-tale eye : how your eyes reveal hidden thoughts and emotions*. New York: Van Nostrand Reinhold Co.

Hilbig, B. E., Thielmann, I., Hepp, J., Klein, S. A., & Zettler, I. (2015). From personality to altruistic behavior (and back): Evidence from a double-blind dictator game. *Journal of Research in Personality, 55*, 46-50. doi:10.1016/j.jrp.2014.12.004

Hilgard, J. (2015, 29 June 2015). Putting PET-PEESE to the Test, Part 1A. *Crystal Prison Zone.* Retrieved from http://crystalprisonzone.blogspot.com/2015/06/putting-pet-peese-to-test-part-1a.html

Hilgard, J. (2016). Functions for R to make PET-PEESE modeling easier. Online: GitHub. Retrieved from https://github.com/Joe-Hilgard/PETPEESE

Hodder, I. (2010). *Religion in the emergence of civilization: Çatalhöyük as a case study*: Cambridge University Press.

Hoffman, G., Forlizzi, J., Ayal, S., Steinfeld, A., Antanitis, J., Hochman, G., . . . Finkenaur, J. (2015). *Robot presence and human honesty: Experimental evidence.* Paper presented at the Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction.

Hoffman, S. (2017). *Managing the State: Social Credit, Surveillance, and the CCP's Plan for China*. Retrieved from https://jamestown.org/program/managing-the-state-social-credit-surveillance-and-the-ccps-plan-for-china/

Hoffman, S. (2018). *Social Credit*. Retrieved from Online: https://www.aspi.org.au/report/social-credit

Hong-Eng, K. (2018, 3 June 2018). Big Brother surveillance? How video cameras can make cities safer and contribute to higher economic growth. *South China Morning Post*. Retrieved from https://www.scmp.com/comment/insight-opinion/article/2148860/big-brother-surveillance-how-video-cameras-can-make-cities

Hoppe, S., Loetscher, T., Morey, S. A., & Bulling, A. (2018). Eye Movements During Everyday Behavior Predict Personality Traits. *Frontiers in Human Neuroscience, 12*(105). doi:10.3389/fnhum.2018.00105

Hossie, T. J., & Sherratt, T. N. (2012). Eyespots interact with body colour to protect caterpillar-like prey from avian predators. *Animal Behaviour, 84*(1), 167-173. doi:10.1016/j.anbehav.2012.04.027

Hossie, T. J., & Sherratt, T. N. (2013). Defensive posture and eyespots deter avian predators from attacking caterpillar models. *Animal Behaviour, 86*(2), 383-389. doi:10.1016/j.anbehav.2013.05.029

Howard, P. N. (2015). *Pax technica : how the internet of things may set us free or lock us up*. New Haven: Yale University Press.

Howard, P. N. (2017). Is Social Media Killing Democracy? Computational Propaganda, Algorithms, Automation & Public Life. [Oxford Internet Institute Inaugural Lecture]. UK: YouTube.

Huang, Z.-w., Liu, L., Zheng, W.-w., Tan, X.-y., & Zhao, X. (2015). Walking the Straight and Narrow: The Moderating Effect of Evaluation Apprehension on the Relationship between Collectivism and Corruption. *PLoS ONE, 10*(3), e0123859. doi:ARTN e0123859

Huber, S., & Huber, O. W. (2012). The centrality of religiosity scale (CRS). *Religions, 3*(3), 710-724. doi:10.3390/rel3030710

Humphrey, N. K. (1976). The social function of intellect. In *Growing points in ethology* (pp. 303-317): Cambridge University Press.

Hutton, R. (1993). *The pagan religions of the ancient British Isles: their nature and legacy*: Wiley-Blackwell.

Imada, T., & Kitayama, S. (2010). Social eyes and choice justification: Culture and dissonance revisited. *Social Cognition, 28*(5), 589-608. doi:10.1521/soco.2010.28.5.589

Ioannidis, J. P. A., & Trikalinos, T. A. (2005). Early extreme contradictory estimates may appear in published research: The Proteus phenomenon in molecular genetics research and randomized trials. *Journal of Clinical Epidemiology, 58*(6), 543-549. doi:https://doi.org/10.1016/j.jclinepi.2004.10.019

Irlenbusch, B., & Ter Meer, J. (2013). Fooling the Nice Guys: Explaining receiver credulity in a public good game with lying and punishment. *Journal of economic behavior & organization, 93*, 321-327. doi:10.1016/j.jebo.2013.03.023

Izuma, K. (2012). The social neuroscience of reputation. *Neuroscience Research, 72*(4), 283-288. doi:http://dx.doi.org/10.1016/j.neures.2012.01.003

Jackson, D. (2015). *Eye donated to charity: The effect of visual monitoring on charitable donations*. Poster. Hanover College.

Johnson, D. (2016). *God is watching you : how the fear of God makes us human*. New York: Oxford University Press.

Johnson, D., & Bering, J. M. (2006). Hand of God, Mind of Man: Punishment and Cognition in the Evolution of Cooperation. *Evolutionary Psychology, 4*(1), 147470490600400119. doi:10.1177/147470490600400119

Johnson, M. H., Griffin, R., Csibra, G., Halit, H., Farroni, T., De Haan, M., . . . Richards, J. (2005). The emergence of the social brain network: Evidence from typical and atypical development. *Development and Psychopathology, 17*(03), 599-619. doi:10.1017/S0954579405050297

Johnson, T., & Wald, C. F. (26 November 2017). The Military Should Teach AI To Watch Drone Footage. *Wired.*

Jolij, J., & de Haan, T. (2014). *Being watched doesn't make you nicer: No effect of visible and invisible eye primes on prosocial behavior in a masked priming study*. Department of Experimental Psychology, Faculty of Behavioural and Social Sciences. Working Paper, Department of Experimental Psychology. University of Groningen. Retrieved from https://pure.rug.nl/ws/files/16082044/JolijDeHaan2014.docx

Jones, D. N., & Paulhus, D. L. (2013). Introducing the Short Dark Triad (SD3): A Brief Measure of Dark Personality Traits. *Assessment, 21*(1), 28-41. doi:10.1177/1073191113514105

Kampe, K. K., Frith, C. D., Dolan, R. J., & Frith, U. (2001). Psychology: Reward value of attractiveness and gaze. *Nature, 413*(6856), 589-589. doi:10.1038/35098149

Kampis, D., Parise, E., Csibra, G., & Kovács, Á. M. (2015). Neural signatures for sustaining object representations attributed to others in preverbal human infants. *Proceedings of the Royal Society of London B: Biological Sciences, 282*(1819). doi:10.1098/rspb.2015.1683

Kandel, E. (2012). *The age of insight: The quest to understand the unconscious in art, mind, and brain, from Vienna 1900 to the present*: Random House.

Kawai, N. (2011). Attentional shift by eye gaze requires joint attention: Eye gaze cues are unique to shift attention. *Japanese Psychological Research, 53*(3), 292-301. doi:10.1111/j.1468-5884.2011.00470.x

Kawashima, R., Sugiura, M., Kato, T., Nakamura, A., Hatano, K., Ito, K., . . . Nakamura, K. (1999). The human amygdala plays an important role in gaze monitoring. *A PET study, 122*(4), 779-783. doi:10.1093/brain/122.4.779

Keep Britain Tidy. (2014). Keeping an Eye on It. Retrieved from http://innovate.keepbritaintidy.org/keeping-an-eye-on-it/2463/2/14/1944/152

Keller, J., & Pfattheicher, S. (2011). Vigilant self-regulation, cues of being watched and cooperativeness. *European Journal of Personality, 25*(5), 363-372. doi:10.1002/per.797

Kilcullen, D. (2013). *Out of the Mountains: The coming age of the urban guerilla*. London: C. Hurst & Co.

Kimbrough, E. O., & Reiss, J. P. (2012). Measuring the Distribution of Spitefulness. *PLoS ONE, 7*(8), e41812. doi:10.1371/journal.pone.0041812

King, Z. M., Henshel, D. S., Flora, L., Cains, M. G., Hoffman, B., & Sample, C. (2018). Characterizing and Measuring Maliciousness for Cybersecurity Risk Assessment. *Frontiers in psychology, 9*(39). doi:10.3389/fpsyg.2018.00039

Kirk, R. E. (1996). Practical Significance: A Concept Whose Time Has Come. *Educational and Psychological Measurement, 56*(5), 746-759. doi:10.1177/0013164496056005002

Kirkpatrick, M. (2010). Facebook's Zuckerberg Says The Age of Privacy is Over. Retrieved from https://readwrite.com/2010/01/09/facebooks_zuckerberg_says_the_age_of_privacy_is_ov/

Kitayama, S., King, A., Yoon, C., Tompson, S., Huff, S., & Liberzon, I. (2014a). The dopamine D4 receptor gene (DRD4) moderates cultural difference in independent versus interdependent social orientation. *Psychological science, 25*(6), 1169-1177. doi:10.1177/0956797614528338

Kitayama, S., King, A., Yoon, C., Tompson, S., Huff, S., & Liberzon, I. (2014b). Survey Measures and Code Book for the Kitayama et al. Psychological Science (in press) paper.

Kitayama, S., Park, H., Sevincer, A. T., Karasawa, M., & Uskul, A. K. (2009). A cultural task analysis of implicit independence: comparing North America, Western Europe, and East Asia. *Journal of personality and social psychology, 97*(2), 236. doi:10.1037/a0015999

Kitayama, S., Snibbe, A. C., Markus, H. R., & Suzuki, T. (2004). Is there any “free” choice? Self and dissonance in two cultures. *Psychological science, 15*(8), 527-533.

Kjernsmo, K., Gronholmi, M., & Merilaita, S. (2016). Adaptive constellations of protective marks: eyespots, eye stripes and diversion of attacks by fish. *Perception, 45*(3), 363-364.

Klein, N., & Epley, N. (2014). The topography of generosity: Asymmetric evaluations of prosocial actions. *Journal of experimental psychology: General, 143*(6), 2366-2379. doi:10.1037/xge0000025

Klein, N., Grossmann, I., Uskul, A. K., Kraus, A. A., & Epley, N. (2015). It pays to be nice, but not really nice: Asymmetric reputations from prosociality across 7 countries. *Judgment and Decision Making, 10*(4), 355.

Knapton, S. (2016). Trees given faces to prevent littering in Forest of Dean. *The Telegraph*. Retrieved from http://www.telegraph.co.uk/science/2016/05/11/trees-given-faces-to-prevent-littering-in-forest-of-dean/

Knowlton, N., & Parker, G. A. (1979). An evolutionarily stable strategy approach to indiscriminate spite. *Nature, 279*, 419. doi:10.1038/279419a0

Kobayashi, H., & Kohshima, S. (1997). Unique morphology of the human eye. *Nature, 387*(6635), 767-768. doi:10.1038/42842

Kobayashi, H., & Kohshima, S. (2001). Unique morphology of the human eye and its adaptive meaning: comparative studies on external morphology of the primate eye. *Journal of Human Evolution, 40*(5), 419-435. doi:https://doi.org/10.1006/jhev.2001.0468

Kobayashi, H., & Kohshima, S. (2011). The gaze that grooms: contribution of social factors to the evolution of primate eye morphology. *Evolution and Human Behavior, 32*(3), 157-165. doi:10.1016/j.evolhumbehav.2010.08.003

Kobayashi, Y., Ohtsuki, H., & Wakano, J. Y. (2016). Population size vs. social connectedness — A gene-culture coevolutionary approach to cumulative cultural evolution. *Theoretical Population Biology, 111*, 87-95. doi:http://dx.doi.org/10.1016/j.tpb.2016.07.001

Krasnow, M. M., & Delton, A. W. (2016). Are Humans Too Generous and Too Punitive? Using Psychological Principles to Further Debates about Human Social Evolution. *Frontiers in psychology, 7*(799). doi:10.3389/fpsyg.2016.00799

Kruskal, W. (1952). A nonparametric analogue based upon ranks of one-way analysis of variance. *Annals of Mathematical Statistics, 23*(1), 140-140.

Kuhn, G., & Kingstone, A. (2009). Look away! Eyes and arrows engage oculomotor responses automatically. *Attention, Perception, & Psychophysics, 71*(2), 314-327. doi:10.3758/APP.71.2.314

Kurzban, R. (2001). The social psychophysics of cooperation: Nonverbal communication in a public goods game. *Journal of Nonverbal Behavior, 25*(4), 241-259. doi: 10.1023/A:1012563421824

Laird, B. K., Bailey, C. D., & Hester, K. (2018). The effects of monitoring environment on problem-solving performance. *The Journal of social psychology, 158*(2), 215-219. doi:10.1080/00224545.2017.1324396

Land, M. F. (2014). *The eye: a very short introduction [electronic resource]*: Oxford : Oxford University Press.

Latané, B., & Nida, S. (1981). Ten years of research on group size and helping. *Psychological bulletin, 89*(2), 308-324. doi:10.1037/0033-2909.89.2.308

Lawson, R. (2004). Small Sample Confidence Intervals for the Odds Ratio. *Communications in Statistics - Simulation and Computation, 33*(4), 1095-1113. doi:10.1081/SAC-200040691

Lee, K., & Ashton, M. C. (2004). Psychometric Properties of the HEXACO Personality Inventory. *Multivariate Behavioral Research, 39*(2), 329-358. doi:10.1207/s15327906mbr3902_8

Lee, T. (2005). Internet control and auto-regulation in Singapore. *Surveillance & Society, 3*(1), 74-95.

Leibbrandt, A., & Lopez-Perez, R. (2011). The Dark Side of Altruistic Third-Party Punishment. *Journal of Conflict Resolution, 55*(5), 761-784. doi:10.1177/0022002711408010

Leiberg, S., Klimecki, O., & Singer, T. (2011). Short-Term Compassion Training Increases Prosocial Behavior in a Newly Developed Prosocial Game (Compassion Training Increases Prosocial Behavior). *PLoS ONE, 6*(3), e17798. doi:10.1371/journal.pone.0017798

Lemoine, J. E., & Roland-Lévy, C. (2017). The effect of the presence of an audience on risk-taking while gambling: the social shield. *Social Influence, 12*(2-3), 101-114. doi:10.1080/15534510.2017.1373697

Levitt, S. D., & List, J. A. (2007a). Viewpoint: On the generalizability of lab behaviour to the field. *Canadian Journal of Economics/Revue canadienne d'économique, 40*(2), 347-370. doi:10.1111/j.1365-2966.2007.00412.x

Levitt, S. D., & List, J. A. (2007b). What Do Laboratory Experiments Measuring Social Preferences Reveal About the Real World? *Journal of Economic Perspectives, 21*(2), 153-174. doi: 10.1257/jep.21.2.153

Levy, J., Foulsham, T., & Kingstone, A. (2013). Monsters are people too. *Biology letters, 9*(1), 20120850. doi:10.1098/rsbl.2012.0850

Li, J., Zhan, Y. L., Fan, W., Liu, L., Li, M., Sun, Y., & Zhong, Y. P. (2018). Sociality Mental Modes Modulate the Processing of Advice-Giving: An Event-Related Potentials Study. *Frontiers in psychology, 9*, 11. doi:10.3389/fpsyg.2018.00042

Limesurvey GmbH. (2003). LimeSurvey: An Open Source survey tool (Version 2.73.1+171220). Hamburg, Germany.: LimeSurvey GmbH,. Retrieved from http://www.limesurvey.org

Lindzon, H. (2010, 10 January 2010). Sorry Zuckerberg/Facebook…The Age of Privacy Ended Long Ago. Retrieved from http://howardlindzon.com/sorry-zuckerbergfacebook-the-age-of-privacy-ended-long-ago/

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, Calif. ; London: Sage Publications.

Liu, J., Li, J., Feng, L., Li, L., Tian, J., & Lee, K. (2014). Seeing Jesus in toast: Neural and behavioral correlates of face pareidolia. *Cortex, 53*, 60-77. doi:https://doi.org/10.1016/j.cortex.2014.01.013

Lönnstedt, O. M., McCormick, M. I., & Chivers, D. P. (2013). Predator-induced changes in the growth of eyes and false eyespots. *3*, 2259. doi:10.1038/srep02259

https://www.nature.com/articles/srep02259#supplementary-information

Luo, Y., Zhang, S., Tao, R., & Geng, H. (2016). The power of subliminal and supraliminal eye contact on social decision making: An individual-difference perspective. *Consciousness and cognition, 40*, 131-140. doi:10.1016/j.concog.2016.01.001

Lyon, D. (2002). Surveillance Studies: understanding visibility, mobility and the phenetic fix. *Surveillance & Society, 1*(1), 1-7.

Lyon, D. (2007). *Surveillance studies: An overview*: Polity.

Lyon, D., Ball, K., & Haggerty, K. D. (2012). *Routledge handbook of surveillance studies*: Routledge.

Lyon, D., & Bauman, Z. (2013). Liquid surveillance: A conversation. *Cambridge, Malden: Polity*.

Macaskill, E., & Dance, G. (2013). *NSA Files: Decoded. What the Revelations Mean for You.* Retrieved from Online: https://www.theguardian.com/world/interactive/2013/nov/01/snowden-nsa-files-surveillance-revelations-decoded#section/1

MacDonald, G., & Leary, M. R. (2005). Why does social exclusion hurt? The relationship between social and physical pain. *Psychological bulletin, 131*(2), 202.

Macrae, C. N., Hood, B. M., Milne, A. B., Rowe, A. C., & Mason, M. F. (2002). Are you looking at me? Eye gaze and person perception. *Psychological science, 13*(5), 460-464.

Manera, V., Elena, M. R., Bayliss, A. P., & Becchio, C. (2014). When Seeing Is More Than Looking: Intentional Gaze Modulates Object Desirability. *Emotion, 14*(4), 824-832. doi:10.1037/a0036258

Mann, S., & Ferenbok, J. (2013). New media and the power politics of sousveillance in a surveillance-dominated world. *Surveillance & Society, 11*(1/2), 18.

Mann, S., Nolan, J., & Wellman, B. (2004). Sousveillance. *Inverse Surveillance in Multimedia Imaging. ACM Multimedia*, 620-627.

Manokha, I. (2018). Surveillance, Panopticism, and Self-Discipline in the Digital Age. *Surveillance & Society, 16*(2), 219-237.

Markus, H. (1978). Effect of Mere Presence on Social Facilitation - Unobtrusive Test. *Journal of Experimental Social Psychology, 14*(4), 389-397. doi:10.1016/0022-1031(78)90034-3

Markus, H. R., & Kitayama, S. (1998). The Cultural Psychology of Personality. *Journal of Cross-Cultural Psychology, 29*(1), 63-87. doi:10.1177/0022022198291004

Marthews, A., & Tucker, C. (2015). Government surveillance and internet search behavior. *Available at SSRN 2412564*.

Maskeliunas, R., & Raudonis, V. (2016). Are you ashamed? Can a gaze tracker tell? *PeerJ Computer Science, 2*, e75. doi:10.7717/peerj-cs.75

Mason, M. F., Tatkow, E. P., & Macrae, C. N. (2005). The look of love gaze shifts and person perception. *Psychological science, 16*(3), 236-239. doi:10.1111/j.0956-7976.2005.00809.x

Mathews, A., Fox, E., Yiend, J., & Calder, A. (2003). The face of fear: Effects of eye gaze and emotion on visual attention. *Visual cognition, 10*(7), 823-835. doi:10.1080/13506280344000095

Matsugasaki, K., Tsukamoto, W., & Ohtsubo, Y. (2015). Two Failed Replications of the Watching Eyes Effect. *Letters on Evolutionary Behavioral Science, 6*(2), 17-20. doi:10.5178/lebs.2015.36

Matsumoto, D. (1999). Culture and self: An empirical assessment of Markus and Kitayama's theory of independent and interdependent self-construals. *Asian Journal of Social Psychology, 2*(3), 289-310.

Maynes, M. J., & Waltner, A. B. (2012). *The family: a world history*. Oxford; New York: Oxford University Press.

Mayr, E. (1963). *Animal species and evolution*. Cambridge, MA: Harvard University Press.

McNees, M. P., Egli, D. S., Marshall, R. S., Schnelle, J. F., & Risley, T. R. (1976). Shoplifting Preventions: Providing InformationThrough Signs. *Journal of Applied Behavior Analysis, 9*(4), 399-405. doi:10.1901/jaba.1976.9-399

Meleady, R., Abrams, D., Van de Vyver, J., Hopthrow, T., Mahmood, L., Player, A., . . . Leite, A. C. (2017). Surveillance or self-surveillance? Behavioral cues can increase the rate of drivers' pro-environmental behavior at a long wait stop. *Environment and Behavior, 49*(10), 1156-1172.

Mifune, N., Hashimoto, H., & Yamagishi, T. (2010). Altruism toward in-group members as a reputation mechanism. *Evolution and Human Behavior, 31*(2), 109-117. doi:10.1016/j.evolhumbehav.2009.09.004

Milgram, S. (1963). Behavioral Study of obedience. *Journal of Abnormal and Social Psychology, 67*(4), 371-378. doi:10.1037/h0040525

Milgram, S. (1974). *Obedience to authority : an experimental view*. London: Tavistock.

Milgram, S. (1977). Subject reaction: The neglected factor in the ethics of experimentation. *Hastings Center Report, 7*(5), 19-23.

Mischel, W., & Shoda, Y. (1995). A cognitive-affective system theory of personality: reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological review, 102*(2), 246.

Monsarrat, P., & Vergnes, J.-N. (2018). The intriguing evolution of effect sizes in biomedical research over time: smaller but more often statistically significant. *GigaScience, 7*(1), gix121-gix121. doi:10.1093/gigascience/gix121

Moonesinghe, R., Khoury, M. J., & Janssens, A. C. J. W. (2007). Most Published Research Findings Are False—But a Little Replication Goes a Long Way. *PLOS Medicine, 4*(2), e28. doi:10.1371/journal.pmed.0040028

Moody, O. (2018). Big Brother is watching them. And we’re next. *The Sunday Times*. Retrieved from https://www.thetimes.co.uk/article/big-brother-is-watching-them-and-we-re-next-858902nbk

Morton, J., & Johnson, M. H. (1991). CONSPEC and CONLERN: a two-process theory of infant face recognition. *Psychological review, 98*(2), 164. doi:10.1037/0033-295X.98.2.164

mos.ru. (28 September 2017). Moscow has one of the world’s largest CCTV systems with face recognition. Retrieved from https://www.mos.ru/en/news/item/30105073/

Mulrooney, M. A., & Ladefoged, T. N. (2005). Hawaiian heiau and agricultural production in the kohala dryland field system. *The Journal of the Polynesian Society, 114*(1), 45-67.

Murphy, R. O., Ackermann, K. A., & Handgraaf, M. (2011). Measuring social value orientation. *Judgment and Decision Making, 6*(8), 771-781.

Na, J., & Kitayama, S. (2012). Will people work hard on a task they choose? Social-eyes priming in different cultural contexts. *Journal of Experimental Social Psychology, 48*(1), 284-290. doi:10.1016/j.jesp.2011.09.003

Nachawati Rego, L. (2012). Syria: The Struggle for Freedom and the End of Silence. *Syria Protests 2011/12*. Retrieved from https://globalvoices.org/2012/01/07/syria-the-struggle-for-freedom-and-the-end-of-silence/

Nelson, C. A. (1997). Electrophysiological correlates of memory development in the first year of life. In H. Reese & M. Franzen (Eds.), *Biological and neuropsychological mechanisms: Life-span developmental psychology* (pp. 95-131). Hillsdale, NJ, US: Lawrence Erlbaum Associates Incorporated.

Nelson, F. (2013). Big Brother (and HMRC) is watching you. *The Spectator: Coffee House Blog*. Retrieved from http://blogs.spectator.co.uk/2013/01/big-brother-and-hmrc-is-watching-you/

Nest. (2017a). Nest Cam IQ: Knows a Person from a Thing. Retrieved from https://nest.com/uk/camera/meet-nest-cam-iq/

Nest. (2017b). Nest.com/uk/about. Retrieved from https://nest.com/uk/about/

Nettle, D., Cronin, K. A., & Bateson, M. (2013). Responses of chimpanzees to cues of conspecific observation. *Animal Behaviour, 86*(3), 595-602. doi:10.1016/j.anbehav.2013.06.015

Nettle, D., Harper, Z., Kidson, A., Stone, R., Penton-Voak, I. S., & Bateson, M. (2013). The watching eyes effect in the Dictator Game: it's not how much you give, it's being seen to give something. *Evolution and Human Behavior, 34*(1), 35-40. doi:http://dx.doi.org/10.1016/j.evolhumbehav.2012.08.004

Nettle, D., Nott, K., & Bateson, M. (2012). ‘Cycle Thieves, We Are Watching You’: Impact of a Simple Signage Intervention against Bicycle Theft (Bicycle Theft). *PLoS ONE, 7*(12), e51738. doi:10.1371/journal.pone.0051738

Neudecker, S. (1989). Eye camouflage and false eyespots: chaetodontid responses to predators. *Environmental Biology of Fishes, 25*(1), 143. doi:10.1007/bf00002208

Nielsen. (2017a). Exelate: A Nielsen Company. Retrieved from http://www.exelate.com

Nielsen. (2017b). Nielsen Marketing Cloud. Retrieved from http://www.nielsen.com/nielsenmarketingcloud

Noah, T., Schul, Y., & Mayo, R. (2018). Thinking of Oneself as an Object of Observation Reduces Reliance on Metacognitive Information. *Journal of Experimental Psychology-General, 147*(7), 1023-1042. doi:10.1037/xge0000440

Noah, T., Schul, Y., & Mayo, R. (2018). When both the original study and its failed replication are correct: Feeling observed eliminates the facial-feedback effect [Press release]

Nock, S. L. (1993). *The costs of privacy: Surveillance and reputation in America*: Aldine de Gruyter.

Norenzayan, A., Shariff, A. F., Gervais, W. M., Willard, A. K., McNamara, R. A., Slingerland, E., & Henrich, J. (2014). The cultural evolution of prosocial religions. *Behavioral and brain sciences, 39*. doi:10.1017/S0140525X14001356

Norris, C., McCahill, M., & Wood, D. (2002). The growth of CCTV: a global perspective on the international diffusion of video surveillance in publicly accessible space. *Surveillance & Society, 2*(2/3).

Northover, S. B., Pedersen, W. C., Cohen, A. B., & Andrews, P. W. (2017a). Artificial surveillance cues do not increase generosity: two meta-analyses. *Evolution and Human Behavior, 38*(1), 144-153. doi:10.1016/j.evolhumbehav.2016.07.001

Northover, S. B., Pedersen, W. C., Cohen, A. B., & Andrews, P. W. (2017b). Effect of artificial surveillance cues on reported moral judgment: Experimental failures to replicate and two meta-analyses. *Evolution and Human Behavior, 38*(5), 561-571. doi:10.1016/j.evolhumbehav.2016.12.003

Oda, R., Kato, Y., & Hiraishi, K. (2015). The Watching-Eye Effect on Prosocial Lying. *Evolutionary Psychology, 13*(3). doi:10.1177/1474704915594959

Oda, R., Niwa, Y., Honma, A., & Hiraishi, K. (2011). An eye- like painting enhances the expectation of a good reputation. *Evolution and Human Behavior, 32*(3), 166-171. doi:10.1016/j.evolhumbehav.2010.11.002

O'Gorman, R., & Silke, A. (2015). Terrorism as altruism: An evolutionary model for understanding terrorist psychology.

Ortmann, A. (2013). On the Interpretation of Giving, Taking, and Destruction in Dictator Games and Joy-of- Destruction Games. In (Vol. 2012-50A): School of Economics, The University of New South Wales.

Palan, S., & Schitter, C. (2018). Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance, 17*, 22-27. doi:https://doi.org/10.1016/j.jbef.2017.12.004

Palermo, R., & Rhodes, G. (2007). Are you always on my mind? A review of how face perception and attention interact. *Neuropsychologia, 45*(1), 75-92. doi:10.1016/j.neuropsychologia.2006.04.025

Palomäki, J., Modic, D., & Yan, J. (2015, August). *Deterring Deception in Online Insurance Claims.* Paper presented at the Decepticon: International Conference on Deceptive Behavior, Cambridge.

Panchanathan, K., & Boyd, R. (2003). A tale of two defectors: the importance of standing for evolution of indirect reciprocity. *Journal of theoretical Biology, 224*(1), 115-126.

Park, J., & Kitayama, S. (2012). Interdependent selves show face-induced facilitation of error processing: cultural neuroscience of self-threat. *Social cognitive and affective neuroscience*, nss125.

Paulhus, D. L., & Williams, K. M. (2002). The Dark Triad of personality: Narcissism, Machiavellianism, and psychopathy. *Journal of Research in Personality, 36*(6), 556-563. doi:https://doi.org/10.1016/S0092-6566(02)00505-6

Penney, J. (2016). Chilling Effects: Online Surveillance and Wikipedia Use. *Berkeley Technology Law Journal.*

Peoples, H. C., & Marlowe, F. W. (2012). Subsistence and the evolution of religion. *Human Nature, 23*(3), 253-269.

Pfattheicher, S., & Keller, J. (2015). The watching eyes phenomenon: The role of a sense of being seen and public self-awareness. *European Journal of Social Psychology, 45*(5), 560-566. doi:10.1002/ejsp.2122

Piazza, J. (2012). “If You Love Me Keep My Commandments”: Religiosity Increases Preference for Rule-Based Moral Arguments. *The International Journal for the Psychology of Religion, 22*(4), 285-302. doi:10.1080/10508619.2011.638598

Piazza, J., Bering, J. M., & Ingram, G. (2011). "Princess Alice is watching you": Children's belief in an invisible person inhibits cheating. *Journal of Experimental Child Psychology, 109*(3), 311-320. doi:10.1016/j.jecp.2011.02.003

Pillutla, M. M., & Murnighan, J. K. (1996). Unfairness, anger, and spite: Emotional rejections of ultimatum offers. *Organizational Behavior and Human Decision Processes, 68*(3), 208-224. doi:10.1006/obhd.1996.0100

Plato. (2003). *The Republic* (D. Lee, Trans.). London: Penguin.

Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of experimental psychology: General, 109*(2), 160-174. doi:10.1037/0096-3445.109.2.160

Powell, K. L., Roberts, G., & Nettle, D. (2012). Eye Images Increase Charitable Donations: Evidence From an Opportunistic Field Experiment in a Supermarket. *ETHOLOGY, 118*(11), 1096-1101. doi:10.1111/eth.12011

Priks, M. (2014). Do Surveillance Cameras Affect Unruly Behavior? A Close Look at Grandstands. *The Scandinavian Journal of Economics, 116*(4), 1160-1179. doi:doi:10.1111/sjoe.12075

Provine, R. R. (2013). The Emotional Eye: Red Sclera as a Uniquely Human Cue of Emotion. *Ethology, 119*(11), 993. doi:10.1111/eth.12144

Provine, R. R., Cabrera, M. O., & Nave-Blodgett, J. (2013). Red, Yellow, and Super-White Sclera. *Human Nature, 24*(2), 126-136. doi:10.1007/s12110-013-9168-x

Provine, R. R., Krosnowski, K. A., & Brocato, N. W. (2009). Tearing: Breakthrough in Human Emotional Signaling. *Evolutionary Psychology, 7*(1), 52-56.

Purzycki, B. G., Apicella, C., Atkinson, Q. D., Cohen, E., McNamara, R. A., Willard, A. K., . . . Henrich, J. (2016). Moralistic gods, supernatural punishment and the expansion of human sociality. *Nature, 530*(7590). doi:10.1038/nature16980

Purzycki, B. G., Apicella, C., Atkinson, Q. D., Cohen, E., McNamara, R. A., Willard, A. K., . . . Henrich, J. (2018). The evolution of religion and morality: a synthesis of ethnographic and experimental evidence from eight societies. *Religion, Brain & Behavior, 8*(2), 101-132. doi:10.1080/2153599X.2016.1267027

R Development Core Team. (2016). R: A language and environment for statistical computing (Version 3.3.1). Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.r-project.org

Raihani, N. J., & Bshary, R. (2012). A positive effect of flowers rather than eye images in a large-scale, cross-cultural dictator game. *Proceedings of the Royal Society B, 279*(1742), 3556-3564. doi:10.1098/rspb.2012.0758

Raudonis, V., Dervinis, G., Vilkauskas, A., Paulauskaite-Taraseviciene, A., & Kersulyte-Raudone, G. (2013). Evaluation of human emotion from eye motions. *Evaluation, 4*(8).

Raymond, N. A., Davies, B. I., Card, B. L., Achkar, Z. A., & Baker, I. L. (2013). While We Watched: Assessing the Impact of the Satellite Sentinel Project. *Georgetown Journal of International Affairs, 14*(2), 185-191.

Reicher, S., & Haslam, S. A. (2006). Rethinking the psychology of tyranny: The BBC prison study. *British Journal of Social Psychology, 45*(1), 1-40. doi:10.1348/014466605X48998

Reicher, S., & Haslam, S. A. (2011). After shock? Towards a social identity explanation of the Milgram ‘obedience’studies. *British Journal of Social Psychology, 50*(1), 163-169.

Reid, A. A., & Andresen, M. A. (2014). An evaluation of CCTV in a car park using police and insurance data. *Security Journal, 27*(1), 55-79. doi:10.1057/sj.2012.14

Reid, V. M., Striano, T., Kaufman, J., & Johnson, M. H. (2004). Eye gaze cueing facilitates neural processing of objects in 4-month-old infants. *Neuroreport, 15*(16), 2553-2555.

Reuters. (2018). China to bar people with bad 'social credit' from planes, trains [Press release]. Retrieved from https://uk.reuters.com/article/uk-china-credit/china-to-bar-people-with-bad-social-credit-from-planes-trains-idUKKCN1GS0ZM

Rigdon, M., Ishii, K., Watabe, M., & Kitayama, S. (2009). Minimal social cues in the dictator game. *Journal of Economic Psychology, 30*(3), 358-367. doi:10.1016/j.joep.2009.02.002

Risse, T., Ropp, S. C., Sikkink, K., Smith, S., Biersteker, T., Brown, C., . . . Groome, J. (1999). *Power of Human Rights*: Cambridge University Press Cambridge.

Rodrigue, C. M. (1992). Can Religion Account for Early Animal Domestications? A Critical Assessment of the Cultural Geographic Argument, Based on Near Eastern Archaeological Data. *The Professional Geographer, 44*(4), 417-430. doi:10.1111/j.0033-0124.1992.00417.x

Roes, F. L., & Raymond, M. (2003). Belief in moralizing gods. *Evolution and Human Behavior, 24*(2), 126-135. doi:https://doi.org/10.1016/S1090-5138(02)00134-4

Rotberg, R. I. (Ed.) (2000). *Peacekeeping and the effective prevention of war*. Washington DC: Brookings Institutions Press.

RowleyConwy, P., Arias, P., Budja, M., Gronenborn, D. Jones, A.,, Louwe Kooijmans, L.,…. & RowelyConwy, P., (2004). How the West was lost was lost: a reconsideration of agricultural origins in Britain, Ireland, and southern Scandinavia. *Current Anthropology*, *45*(S4), S83-S113. doi: 10.1086/422083

Rozin, P., & Royzman, E. B. (2001). Negativity Bias, Negativity Dominance, and Contagion. *Personality & Social Psychology Review (Lawrence Erlbaum Associates), 5*(4), 296-320. doi:https://doi.org/10.1207/S15327957PSPR0504_2

Ruxton, G. D., & Neuhäuser, M. (2012). Review of alternative approaches to calculation of a confidence interval for the odds ratio of a 2 × 2 contingency table. *Methods in ecology and evolution, 4*(1), 9-13. doi:10.1111/j.2041-210x.2012.00250.x

Ryan, R., & Deci, E. (2000). Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. *Am Psychol, 55*(1), 68-78.

Sadrieh, A., & Schröder, M. (2012). The Desire to Influence Others. *IDEAS Working Paper Series from RePEc*.

Safina, C. (2016). Keeping peace with predators can cut livestock deaths.

Saiidi, U. (2018). Singapore's crime rate is so low that many shops don't even lock up. *CNBC*. Retrieved from cnbc.com website: https://www.cnbc.com/2018/01/16/singapores-crime-rate-is-so-low-that-many-shops-dont-even-lock-up.html

Sander, D., Grandjean, D., Kaiser, S., Wehrle, T., & Scherer, K. R. (2007). Interaction effects of perceived gaze direction and dynamic facial expression: Evidence for appraisal theories of emotion. *European Journal of Cognitive Psychology, 19*(3), 470-480. doi:10.1080/09541440600757426

Sandhu, A. (2016). Camera-friendly policing: how the police respond to cameras and photographers. *Surveillance & Society, 14*(1), 78.

Sandvik, K., & Raymond, N. (2017). Beyond the Protective Effect: Towards a Theory of Harm for Information Communication Technologies in Mass Atrocity Response. *Genocide Studies and Prevention: An International Journal,, 11*(1).

Sartre, J.-P. (1978). *Being and Nothingness: An Essay on Phenomenological Ontology* (H. E. Barnes, Trans.). Abingdon: Routledge.

Satellite Sentinel Project. (2010). George Clooney: 'The World is Watching' – New Satellite Sentinel Project Aims to Deter War in Sudan [Press release]. Retrieved from http://www.satsentinel.org/press-release/george-clooney-'-world-watching'-—-new-satellite-sentinel-project-aims-deter-war-sudan

Sato, W., Yoshikawa, S., Kochiyama, T., & Matsumura, M. (2004). The amygdala processes the emotional significance of facial expressions: an fMRI investigation using the interaction between expression and face direction. *NeuroImage, 22*(2), 1006-1013. doi:10.1016/j.neuroimage.2004.02.030

Scaife, M. (1976a). The response to eye-like shapes by birds II. The importance of staring, pairedness and shape. *Animal Behaviour, 24*(1), 200-206. doi:10.1016/S0003-3472(76)80116-9

Scaife, M. (1976b). The response to eye-like shapes by birds. I. The effect of context: A predator and a strange bird. *Animal Behaviour, 24*(1), 195-199. doi:10.1016/S0003-3472(76)80115-7

Scheier, M. F., & Carver, C. S. (1985). The Self-Consciousness Scale: A Revised Version for Use with General Populations. *Journal of Applied Social Psychology, 15*(8), 687-699.

Schilbach, L. (2015). Eye to Eye, face to face and brain to brain: novel approaches to study behavioral dynamics and neural mechanisms of social interactions. *Current Opinion in Behavioral Sciences*(3), 130-135. doi:10.1016/j.cobeha.2015.03.006

Schilbach, L., Eickhoff, S., Cieslik, E., Shah, N., Fink, G., & Vogeley, K. (2011). Eyes on me: an fMRI study of the effects of social gaze on action control. *Social Cognition and Affective Neuroscience, 6*(4), 393-403. doi:10.1093/scan/nsq067

Schjoedt, U., Stødkilde-Jørgensen, H., Geertz, A. W., & Roepstorff, A. (2009). Highly religious participants recruit areas of social cognition in personal prayer. *Social cognitive and affective neuroscience, 4*(2), 199-207. doi:10.1093/scan/nsn050

Scott-Phillips, T. C., Dickins, T. E., & West, S. A. (2011). Evolutionary Theory and the Ultimate–Proximate Distinction in the Human Behavioral Sciences. *Perspectives on Psychological Science, 6*(1), 38-47. doi:10.1177/1745691610393528

Schmitt, B. H., Gilovich, T., Goore, N., & Joseph, L. (1986). Mere presence and social facilitation: One more time. *Journal of Experimental Social Psychology, 22*(3), 242-248. doi:https://doi.org/10.1016/0022-1031(86)90027-2

Senju, A., & Hasegawa, T. (2006). Do the upright eyes have it? *Psychonomic bulletin & review, 13*(2), 223-228. doi:10.3758/BF03193834

Senju, A., & Johnson, M. H. (2009). The eye contact effect: mechanisms and development. *Trends in Cognitive Sciences, 13*(3), 127-134. doi:10.1016/j.tics.2008.11.009

Serb, J. M., & Eernisse, D. J. (2008). Charting Evolution's Trajectory: Using Molluscan Eye Diversity to Understand Parallel and Convergent Evolution. *Evolution: Education and Outreach, 1*(4), 439-447. doi:10.1007/s12052-008-0084-1

Shanley, M. (2012, 2 October 2012). Live from Syria: Swedish streaming site gains followers. *Reuters*. Retrieved from https://www.reuters.com/article/us-sweden-bambuser/live-from-syria-swedish-streaming-site-gains-followers-idUSBRE8910D620121002

Shariff, A. F., & Norenzayan, A. (2007). God is watching you: Priming God concepts increases prosocial behavior in an anonymous economic game. *Psychological science, 18*(9), 803-809. doi:DOI 10.1111/j.1467-9280.2007.01983.x

Shariff, A. F., Willard, A. K., Andersen, T., & Norenzayan, A. (2015). Religious priming a meta-analysis with a focus on prosociality. *Personality and Social Psychology Review, 20*(1), 1088868314568811. doi:10.1177/1088868314568811

Shirado, H., & Christakis, N. A. (2017). Locally noisy autonomous agents improve global human coordination in network experiments. *Nature, 545*, 370. doi:10.1038/nature22332

https://www.nature.com/articles/nature22332#supplementary-information

Sim, R. (19 March 2016). More surveillance cameras as deterrent. *Straits Times*.

Singal, J. (2018). Updated: A re-replication of a psychological classic provides a cautionary tale about overhyped science.  Retrieved from https://digest.bps.org.uk/2018/08/15/a-re-replication-of-a-psychological-classic-provides-a-cautionary-tale-about-overhyped-science/

Singelis, T. M. (1994). The measurement of independent and interdependent self-construals. *Personality and Social Psychology Bulletin, 20*(5), 580-591.

Sinha, A. G., Neier, A., McFarland Sanchez-Moreno, M., Prasow, A., Resiser, S., Smith, P., . . . Young, P. (2014). *With Liberty to Monitor All: How Large-Scale US Surveillance is Harming Journalism, Law, and American Democracy*. Retrieved from online: https://www.hrw.org/report/2014/07/28/liberty-monitor-all/how-large-scale-us-surveillance-harming-journalism-law-and

Sisk, T. (2017). Peacekeeping in Peril: The First 65 Years. *International Peacekeeping, 24*(3), 493-499. doi:10.1080/13533312.2017.1317074

Sivarajasingam, V., Shepherd, J. P., & Matthews, K. (2003). Effect of urban closed circuit television on assault injury and violence detection. *Injury Prevention, 9*(4), 312-316. doi:10.1136/ip.9.4.312

Snowden, E. (23 January 2014). Snowden Q&A: Indiscriminate mass surveillance is the biggest problem we face. Retrieved from https://www.rt.com/usa/snowden-obama-whistleblower-protection-101/

Snyder, M. (1974). Self-monitoring of expressive behavior. *Journal of personality and social psychology, 30*(4), 526. doi: 10.1037/h0037039

Sparks, A., & Barclay, P. (2013). Eye images increase generosity, but not for long: the limited effect of a false cue. *Evolution and Human Behavior, 34*(5), 317-322. doi:10.1016/j.evolhumbehav.2013.05.001

Stanley, M. (1974). Obedience to authority. *An Experimental View. Harper, New York.*

Starr, S. (2012). *Revolt in Syria: Eye-witness to the Uprising*: Hurst.

Statista. (2017). Number of monthly active Facebook users worldwide as of 1st quarter 2017 (in millions). *Statista: The Statistics Portal: Data and Studies from more than 18,000 sources.* Retrieved from https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/

Steinmetz, J., Xu, Q., Fishbach, A., & Zhang, Y. (2016). Being Observed Magnifies Action. *Journal of personality and social psychology, 111*(6), 852-865. doi:10.1037/pspi0000065

Strack, F. (2016). Reflection on the smiling registered replication report [Press release]

Strack, F., Martin, L. L., & Stepper, S. (1988). Inhibiting and facilitating conditions of the human smile: A nonobtrusive test of the facial feedback hypothesis. *Journal of personality and social psychology, 54*(5), 768-777. doi:10.1037/0022-3514.54.5.768

Strait, D., Grine, F. E., & Fleagle, J. G. (2015). Analyzing Hominin Hominin Phylogeny: Cladistic Approach. In W. Henke & I. Tattersall (Eds.), *Handbook of Paleoanthropology* (pp. 1989-2014). Berlin, Heidelberg: Springer.

Swanson, G. E. (1964). *The birth of the gods: The origin of primitive beliefs* (Vol. 93): University of Michigan Press.

Tanaka, H., Ohtsuki, H., & Ohtsubo, Y. (2016). The price of being seen to be just: an intention signalling strategy for indirect reciprocity. *Proceedings of the Royal Society B: Biological Sciences, 283*(1835). doi:10.1098/rspb.2016.0694

Tane, K., & Takezawa, M. (2011). Perception of Human Face Does Not Induce Cooperation in Darkness. *Letters on Evolutionary Behavioural Science, 2*(2), 24-27. doi:1884-927X

TechCrunch (Writer). (2010). Facebook CEO Mark Zuckerberg TechCrunch Interview with TechCrunch Founder Michael Arrington at the Crunchies [YouTube Video]. In *TechCrunch, Friday, January 8th, San Francisco*. YouTube: Zennie62 Live Vlog.

The Behavioural Insights Team. (2017). *The Behavioural Insights Team Update Report 2016-17*. Retrieved from London: https://www.behaviouralinsights.co.uk/publications/the-behavioural-insights-team-update-report-2016-17/

Tinbergen, N. (1963). On aims and methods of Ethology. *Zeitschrift für Tierpsychologie, 20*(4), 410-433. doi:doi:10.1111/j.1439-0310.1963.tb01161.x

Tomasello, M., Hare, B., Lehmann, H., & Call, J. (2007). Reliance on head versus eyes in the gaze following of great apes and human infants: the cooperative eye hypothesis. *Journal of Human Evolution, 52*(3), 314-320. doi:10.1016/j.jhevol.2006.10.001

Tomasello, M. (2009). *The cultural origins of human cognition*: Harvard university press.

Tomasello, M., Melis, A. P., Tennie, C., Wyman, E., & Herrmann, E. (2012). Two Key Steps in the Evolution of Human Cooperation: The Interdependence Hypothesis. *Current Anthropology, 53*(6), 673-692. doi:10.1086/668207

Tomasello, M., & Vaish, A. (2013). Origins of Human Cooperation and Morality. *Annual review of psychology, 64*(1), 231-255. doi:10.1146/annurev-psych-113011-143812

Tooby, J., & Cosmides, L. (1997). Foreword. In S. Baron-Cohen (Ed.), *Mindblindness: An essay on autism and theory of mind*: MIT press.

Topál, J., & Csányi, V. (1994). The effect of eye- like schema on shuttling activity of wild house mice ( Mus musculus domesticus): Context- dependent threatening aspects of the eyespot patterns. *Animal Learning & Behavior, 22*(1), 96-102. doi:10.3758/BF03199961

towerd@ta. (2017). Enhance your Email List with Email Intelligence. Retrieved from http://www.towerdata.com/email-intelligence/email-enhancement

Traver, C., & Cordell, T. (2014). *Look Me in the Eyes: Examining the Effects of Priming with Eyes on Deception*. Retrieved from http://christophertraver.com/assets/eye-primes.pdf

Triandis, H. C. (1989). The self and social behavior in differing cultural contexts. *Psychological review, 96*(3), 506. doi: 10.1037/0033-295x.96.3.506

Triplett, N. (1898). The Dynamogenic Factors in Pacemaking and Competition. *The American journal of psychology, 9*(4), 507-533. doi:10.2307/1412188

Trivers, R. L. (1971). The Evolution of Reciprocal Altruism. *The Quarterly Review of Biology, 46*(1), 35-57.

Tsao, D. Y., Freiwald, W. A., Tootell, R. B. H., & Livingstone, M. S. (2006). A cortical region consisting entirely of face-selective cells. *Science, 311*(5761), 670-674. doi:10.1126/science.1119983

Tsao, D. Y., Moeller, S., & Freiwald, W. A. (2008). Comparing face patch systems in macaques and humans. *Proceedings of the national Academy of Sciences, 105*(49), 19514-19519. doi:10.1073/pnas.0809662105

Tullet, A. (18 August 2018). A Not-so-Hypothetical HIBAR. Retrieved from http://sometimesimwrong.typepad.com/wrong/2018/08/hibar.html

Charter of the United Nations Chapter VI: Pacific Settlement of Disputes, (24 October 1945).

Vaish, A., Kelsey, C. M., Tripathi, A., & Grossmann, T. (2017). Attentiveness to eyes predicts generosity in a reputation-relevant context. *Evolution and Human Behavior, 38*(6), 729-733. doi:10.1016/j.evolhumbehav.2017.07.004

van Bommel, M., van Prooijen, J. W., Elffers, H., & Van Lange, P. A. M. (2012). Be aware to care: Public self-awareness leads to a reversal of the bystander effect. *Journal of Experimental Social Psychology, 48*(4), 926-930. doi:10.1016/j.jesp.2012.02.011

van Bommel, M., van Prooijen, J. W., Elffers, H., & van Lange, P. A. M. (2014). Intervene to be Seen: The Power of a Camera in Attenuating the Bystander Effect. *Social Psychological and Personality Science, 5*(4), 459-466. doi:10.1177/1948550613507958

Van Lange, P. A. (1999). The pursuit of joint outcomes and equality in outcomes: An integrative model of social value orientation. *Journal of personality and social psychology, 77*(2), 337.

Van Vugt, M., Roberts, G., & Hardy, C. (2007). Competitive altruism: Development of reputation-based cooperation in groups. In R. Dunbar & L. Barrett (Eds.), *Handbook of evolutionary psychology* (pp. 531-540). Oxford Handbooks Online: Oxford University Press.

Vogt, S., Efferson, C., Berger, J., & Fehr, E. (2015). Eye spots do not increase altruism in children. *Evolution and Human Behavior, 36*(3), 224-231. doi:10.1016/j.evolhumbehav.2014.11.007

Vanlange, P., Otten, W., Debruin, E., & Joireman, J. (1997). Development of prosocial, individualistic, and competitive orientations: Theory and preliminary evidence. *J. Pers. Soc. Psychol., 73*(4), 733-746. doi: 10.1037/0022-3514.73.6.1330

von Grünau, M., & Anston, C. (1995). The detection of gaze direction: A stare-in-the-crowd effect. *Perception, 24*(11), 1297-1313. doi:10.1068/p241297

Voors, M., Turley, T., Kontoleon, A., Bulte, E., & List, J. A. (2012). Exploring whether behavior in context-free experiments is predictive of behavior in the field: Evidence from lab and field experiments in rural Sierra Leone. *Economics Letters, 114*(3), 308-311. doi:https://doi.org/10.1016/j.econlet.2011.10.016

Vuilleumier, P., George, N., Lister, V., Armony, J., & Driver, J. (2005). Effects of perceived mutual gaze and gender on face processing and recognition memory. *Visual cognition, 12*(1), 85-101. doi:10.1080/13506280444000120

Waddel, K. (2016). The Privacy Problem with Digital Assistants. *The Atlantic*. Retrieved from https://www.theatlantic.com/technology/archive/2016/05/the-privacy-problem-with-digital-assistants/483950/

Wagenmakers, E.-J., Beek, T., Dijkhoff, L., Gronau, Q. F., Acosta, A., Adams, R. B., . . . Zwaan, R. A. (2016). Registered Replication Report:Strack, Martin, & Stepper (1988). *Perspectives on Psychological Science, 11*(6), 917-928. doi:10.1177/1745691616674458

Wald, C. F., & Johnson, T. (13 June 2017). Military Omnipresence: A Unifying Concept for America's 21st-Century Fighting Edge. Retrieved from http://www.defenseone.com/ideas/2017/06/military-omnipresence-unifying-concept-americas-21st-century-fighting-edge/138640/

Wang, Y., & Hamilton, A. F. d. C. (2014). Why does gaze enhance mimicry? Placing gaze-mimicry effects in relation to other gaze phenomena. *The Quarterly Journal of Experimental Psychology, 67*(4), 747-762. doi:10.1080/17470218.2013.828316

Wang, Y., Ramsey, R., & Hamilton, A. (2011). The control of mimicry by eye contact is mediated by medial prefrontal control. *Journal of Neuroscience, 31*(33), 12001-12010. doi:10.1523/Jneurosci.0845-11.2011

Warrington, E. K., & Shallice, T. (1984). Category specific semantic impairments. *Brain, 107*(3), 829-853. doi:10.1093/brain/107.3.829

Waters, N. (2018, 25 September 2018). Google Maps Is a Better Spy Than James Bond. *Foreign Policy Magazine*.

Watson-Jones, R. E., & Legare, C. H. (2016). The Social Functions of Group Rituals. *Current Directions in Psychological Science, 25*(1), 42-46. doi:10.1177/0963721415618486

Webber, M. (2017). How many CCTV cameras are installed in Singapore? Retrieved from https://www.quora.com/How-many-CCTV-cameras-are-installed-in-Singapore

Weisburd, D., Telep, C. W., Hinkle, J. C., & Eck, J. E. (2010). Is problem-oriented policing effective in reducing crime and disorder? *Criminology & Public Policy, 9*(1), 139-172. doi:10.1111/j.1745-9133.2010.00617.x

Welsh, B., & Farrington, D. (2008). Effects of improved street lighting on crime: a systematic review. *Campbell systematic reviews, 4*(13).

Welsh, B., & Farrington, D. (2009). Public Area CCTV and Crime Prevention: An Updated Systematic Review and Meta- Analysis. *Justice Quarterly, 26*(4), 716-745. doi:10.1080/07418820802506206

West Midlands Police. (2006). W Mids Gain 'Momentum' in Fight Against Crime [Press release]. Retrieved from http://www.policeoracle.com/news/W-Mids-Gain-Momentum-In-Fight-Against-Crime_11284.html

Whetham, D. (2015). Drones to protect. *The International Journal of Human Rights, 19*(2), 199-210. doi:10.1080/13642987.2014.991215

White, C. (2015). *Unexpected events as a cue to social surveillance.* (BSc), King’s University College, Western University, London, Ontario, Canada, Online. Retrieved from http://ir.lib.uwo.ca/cgi/viewcontent.cgi?article=1018&context=psychK_uht

White, L. A. (1959). *The Evolution of Culture: The Development of Civilization to the Fall of Rome* (2007 ed.). New York: Left Coast Press.

White, R. W. (1959). Motivation reconsidered: The concept of competence. *Psychological review, 66*(5), 297-333.

Wickler, W. (1968). Mimicry in plants and animals.

Wickler, W. (2013). Understanding Mimicry – with Special Reference to Vocal Mimicry. *ETHOLOGY, 119*(4), 259-269. doi:doi:10.1111/eth.12061

Wicklund, R., & Duval, S. (1971). Opinion change and performance facilitation as a result of objective self-awareness. *Journal of Experimental Social Psychology, 7*(3), 319-342. doi:http://dx.doi.org/10.1016/0022-1031(71)90032-1

Wicklund, R., & Duval, S. (1972). A theory of objective self awareness. In: Oxford, England: Academic Press.

Wiekens, C. J., & Stapel, D. A. (2010). Self-Awareness and Saliency of Social Versus Individualistic Behavioral Standards. *Social Psychology, 41*(1), 10-19. doi:10.1027/1864-9335/a000003

Wikipedia. (2018). Syrian Observatory for Human Rights. *Wikipedia.* Retrieved from https://en.wikipedia.org/wiki/Syrian_Observatory_for_Human_Rights

Wilcox, R. (2017). Chapter 10 - Robust Regression. In R. Wilcox (Ed.), *Introduction to Robust Estimation and Hypothesis Testing (Fourth Edition)* (pp. 517-583): Academic Press.

Willis, C., Marshall, P., McKinley, J., Pitts, M., Pollard, J., Richards, C., . . . Pearson, M. P. (2016). The dead of Stonehenge. *Antiquity, 90*(350), 337-356. doi:10.15184/aqy.2016.26

Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after 100-ms exposure to a face. *Psychological Science*, *17*, 592–598. doi:10.1111/j.1467-9280.2006.01750.x

Wilson, D. S. (1975). A theory of group selection. *Proceedings of the national Academy of Sciences, 72*(1), 143-146. doi:10.1073/pnas.72.1.143

Wilson, D. S., Wilczynski, C., Wells, A., & Weiser, L. (2000). The evolution of cognition.

Wilson, T. D., Reinhard, D. A., Westgate, E. C., Gilbert, D. T., Ellerbeck, N., Hahn, C., . . . Shaked, A. (2014). Just think: The challenges of the disengaged mind. *Science, 345*(6192), 75-77. doi:10.1126/science.1250830

Wohlsen, M. (2014). What Google Really Gets out of Buying Nest of $3.2 Billion. *Wired*. Retrieved from: https://www.wired.com/2014/01/googles-3-billion-nest-buy-finally-make-internet-things-real-us/

Wojciszke, B., & Abele, A. E. (2008). The primacy of communion over agency and its reversals in evaluations. *European Journal of Social Psychology*, *38*, 1139–1147. doi:10.1002/ejsp.549

Yarbus, A. L. (1967). Eye movements during perception of complex objects. In A. L. Yarbus (Ed.), *Eye Movements and Vision* (pp. 171-211). USSR: Springer US.

Yoeli, E., Hoffman, M., Rand, D. G., & Nowak, M. A. (2013). Powering up with indirect reciprocity in a large-scale field experiment. *Proceedings of the national Academy of Sciences, 110*(Supplement 2), 10424-10429. doi:10.1073/pnas.1301210110

Yoshie, M., Nagai, Y., Critchley, H. D., & Harrison, N. A. (2016). Why I tense up when you watch me: Inferior parietal cortex mediates an audience's influence on motor performance. *Scientific reports, 6*. ARTN 19305 10.1038/srep19305

Zajonc, R. B. (1965). *Social facilitation*: Research Center for Group Dynamics, Institute for Social Research, University of Michigan.

Zajonc, R. B., Heingart, A., & Herman, E. M. (1969). Social Enhancement and Impairment of Performance in the Cockroach. *Journal of personality and social psychology, 13*(2), 83-92. doi:10.1037/h0028063

Zeitner, R. M., & Weight, D. G. (1979). The pupillometric response as a parameter of self-esteem. *Journal of Clinical Psychology, 35*(1), 176-183. doi:doi:10.1002/1097-4679(197901)35:1<176::AID-JCLP2270350128>3.0.CO;2-Y

Zengerink, L. (2013). *The image of eyes: psychological mechanisms and effects on littering in a real-world setting.* (MSc), University of Twente, Online. Retrieved from http://purl.utwente.nl/essays/64617

Zhong, C.-B., Bohns, V. K., & Gino, F. (2010). Good lamps are the best police: darkness increases dishonesty and self-interested behavior. *Psychological science, 21*(3), 311-314. doi:10.1177/0956797609360754

Zimbardo, P. (2006). On rethinking the psychology of tyranny: the BBC prison study. *The British Journal of Social Psychology, 45*, 47-53.

Zimbardo, P. (2011). *The Lucifer effect: How good people turn evil*: Random House.

Zuboff, S. (2015). Big other: surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology, 30*(1), 75-89. doi:10.1057/jit.2015.5

Zuckerman, E. (2016). Why we must continue to turn the camera on Police. *MIT Technology Review*. Retrieved from https://www.technologyreview.com/s/601878/why-we-must-continue-to-turn-the-camera-on-police/

Zuo, S., Huang, N., Cai, P., & Wang, F. (2018). The lure of antagonistic social strategy in unstable socioecological environment: Residential mobility facilitates individuals' antisocial behavior. *Evolution and Human Behavior, 39*(3), 364-371. doi:10.1016/j.evolhumbehav.2018.03.002

# Appendices

# A

# Appendix A. Chapter 2, Study 1, Perceived Surveillance & Prosociality

---

## Additional Analyses

In addition to the exploratory variables reported in Chapter 2, data was also collected to try to identify whether participants religiosity might predict interact with surveillance conditions to predict their responsiveness to surveillance cues. A two-way independent factorial ANOVA was used to look for interaction effects between Surveillance Condition, Religion and Helping rates, finding no significant interaction effect between condition and religion on helping rates $F(3, 191)=0.91, p=0.44$. A two-way independent factorial ANOVA examining Surveillance Condition and Religion (religious/not religious) effects on donations in the Dictator game was also non-significant, $F(3, 191)=1.14, p=0.34$. This analysis was not included in the main results section as the data collected was in response to the overly-simple binary question: ‘Are you religious or not?’. The question was too simple but the answers – which were free text – were too complex. To make any sense of our data, we had to amalgamate multiple answers into categories to provide something intelligible, in the end deciding that the only way to do this sensibly was to create a binary list of those who were religious and those who were not – to draw the line anywhere else was almost impossible, and resulted in far too many arbitrary categories. However, creating a binary outcome measure was still unrepresentative. It required subsuming data from too broad a range to representative of the actual answers received. Responses ranged from “ religious but not practicing”, “agnostic but not religious”, “spiritual but not religious” and even a few “Jedis”. And that’s just those that are reportable. When testing military participants, it may be better not to provide free-text answer opportunities.

It would have been better to have used a Lickert-type 7-point scale to increase the sensitivity of our analysis, or to have gone further and used a validated measure to test for an individual’s religiosity, such as the centrality of religiosity scale (Huber & Huber, 2012).

Furthermore, our dataset was also insufficiently ethnically diverse to draw conclusions on the effect of ethnicity [additionally, an ANOVA found no interaction effects between ethnicity and responsiveness to any of the four surveillance cues $F(23,169)=1.059$, $p=0.39$].

# B

# Appendix B for Chapter 4

---

## Tables & Charts from Joy of Destruction Game Analyses

### *Public Self-Consciousness Scores Analysis for Studies 1-3 & all 3 Studies in the Aggregate.*

Table A1. Study 1. Poisson GLM Examining Destruction rates in the JoD by Condition, Public Self-Consciousness

| | *B* | *SE B* | *Z* | *P* |
|---|---|---|---|---|
| Constant | 0.63 | 0.20 | 3.13 | <0.001 *** |
| Condition: Eyes | 0.16 | 0.28 | 0.57 | 0.57 |
| Condition: Camera | -0.64 | 0.31 | -2.09 | 0.04* |
| Public Self-Consciousness | 0.02 | 0.02 | 1.23 | 0.22 |
| Condition Eyes: Public Self-Consciousness | -0.01 | 0.02 | -0.36 | 0.72 |
| Condition Camera: Public Self- Consciousness | 0.05 | 0.02 | 1.84 | 0.07 . |

Table A2. Study 2. Poisson GLM Examining Destruction rates in the JoD by Condition, Public Self-Consciousness

| | *B* | *SE B* | *Z* | *P* |
|---|---|---|---|---|
| Constant | 1.38 | 0.13 | 10.98 | <0.001*** |
| Condition: Eyes | -0.56 | 0.19 | -2.93 | <0.01** |
| Condition: Camera | -0.63 | 0.21 | -2.95 | <0.01** |
| Public Self-Consciousness | -0.04 | 0.02 | 3.09 | <0.01** |
| Condition Eyes: Public Self-Consciousness | 0.06 | 0.02 | 3.09 | <0.01** |
| Condition Camera: Public Self- Consciousness | 0.04 | 0.02 | 2.35 | 0.02* |

Table A3. Study 3. Poisson GLM Examining Destruction rates in the JoD by Condition, Public Self-Consciousness

| | *B* | *SE B* | *Z* | *P* |
|---|---|---|---|---|
| Constant | 0.72 | 0.23 | 3.21 | 0.001** |
| Condition: Eyes | 0.08 | 0.31 | 0.25 | 0.80 |
| Condition: Camera | -0.06 | 0.38 | -0.17 | 0.86 |
| Public Self-Consciousness | 0.02 | 0.02 | 1.25 | 0.21 |
| Condition Eyes: Public Self-Consciousness | -0.02 | 0.03 | -0.78 | 0.44 |
| Condition Camera: Public Self- Consciousness | -0.02 | 0.03 | -0.78 | 0.44 |

Table A4. Aggregate Analysis. Poisson GLM Examining Destruction rates in the JoD by Condition, Public Self-Consciousness

| | *B* | *SE B* | *Z* | *P* |
|---|---|---|---|---|
| Constant | 1.21 | 0.09 | 13.85 | <0.001 *** |
| Condition: Eyes | -0.30 | 0.13 | -2.37 | 0.02 * |
| Condition: Camera | -0.36 | 0.14 | -2.51 | 0.01 * |
| Public Self-Consciousness | -0.02 | 0.01 | -3.13 | 0.002 ** |
| Condition Eyes: Public Self-Consciousness | 0.02 | 0.01 | 2.01 | 0.04* |
| Condition Camera: Public Self- Consciousness | 0.01 | 0.01 | 1.01 | 0.31 |

*Figure A1. Study 1. The influence of Public Self-Consciousness (PubSC) on Destruction Rates in the Joy of Destruction Game.*

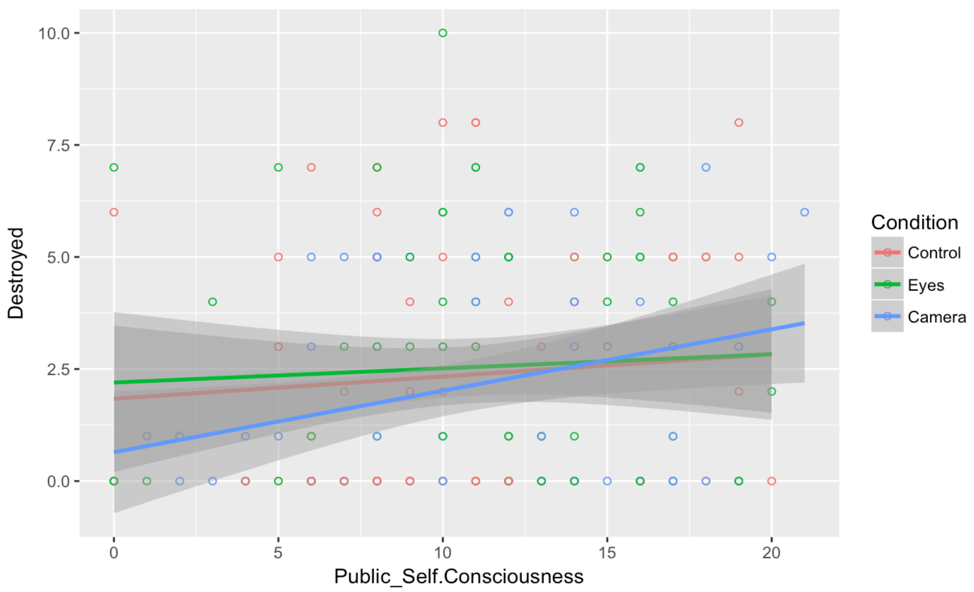


*Figure A2. Study 2. The influence of Public Self-Consciousness (PubSC) on Destruction Rates in the Joy of Destruction Game.*

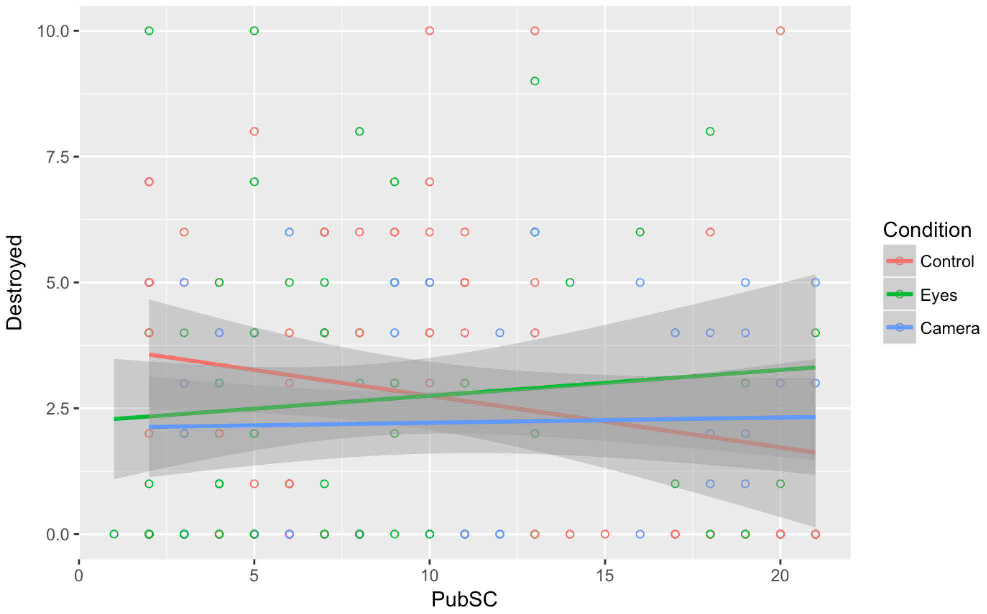


*Figure A3. Study 3. The influence of Public Self-Consciousness (PubSC) on Destruction Rates in the Joy of Destruction Game.*

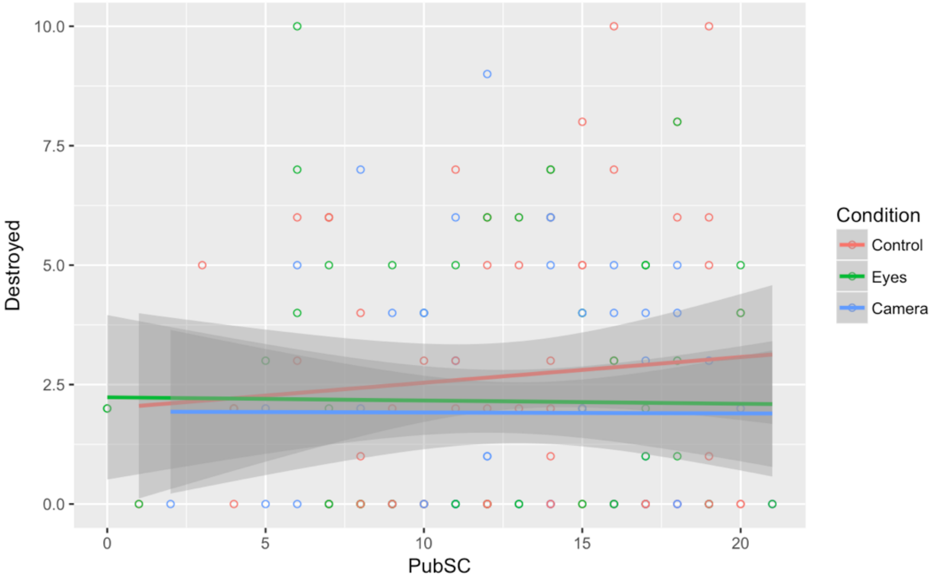
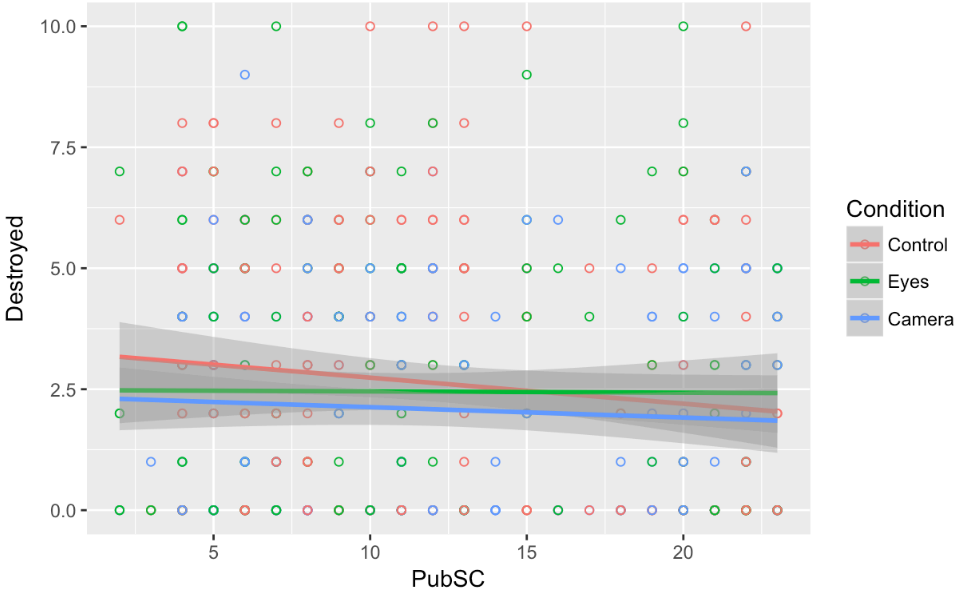


*Figure A4. Aggregate Analysis. The influence of Public Self-Consciousness (PubSC) on Destruction Rates in the Joy of Destruction Game.*



## *Interdependence Scores Analysis for Studies 1-3 & all 3 Studies in the Aggregate.*

Table A5. Study 1. Poisson GLM Examining Destruction rates in the JoD by Condition, Independence-Interdependent Self-Construal

| | *B* | *SE B* | *Z* | *P* |
|---|---|---|---|---|
| Constant | 0.83 | 0.08 | 10.29 | <.001 *** |
| Condition: Eyes | 0.05 | 0.11 | 0.46 | .64 |
| Condition: Camera | -0.09 | 0.12 | -0.77 | .44 |
| Independence-Interdependence | 0.18 | 0.08 | 2.36 | .02* |
| Condition Eyes: Independence-Interdependence | 0.08 | 0.12 | 0.69 | .49 |
| Condition Camera: Independence-Interdependence | -0.45 | 0.13 | -3.63 | <.001 *** |

Table A6. Study 2. Poisson GLM Examining Destruction rates in the JoD by Condition, Independence-Interdependent Self-Construal

| | *B* | *SE B* | *Z* | *P* |
|---|---|---|---|---|
| Constant | 1.03 | 0.07 | 14.45 | <.001*** |
| Condition: Eyes | -0.05 | 0.10 | -0.48 | .63 |
| Condition: Camera | -0.22 | 0.12 | -1.87 | .06 . |
| Independence-Interdependence | 0.002 | 0.09 | 0.02 | .98 |
| Condition Eyes: Independence-Interdependence | 0.17 | 0.12 | 1.49 | .14 |
| Condition Camera: Independence-Interdependence | 0.13 | 0.14 | 0.93 | .35 |

Table A7. Study 3. Poisson GLM Examining Destruction rates in the JoD by Condition, Independent-Interdependent Self-Construal

| | *B* | *SE B* | *Z* | *P* |
|---|---|---|---|---|
| Constant | 0.99 | 0.07 | 13.61 | <.001*** |
| Condition: Eyes | -0.22 | 0.11 | -1.97 | .05* |
| Condition: Camera | -0.46 | 0.14 | -3.28 | .001** |
| Independence-Interdependence | 0.07 | 0.07 | 1.01 | .31 |
| Condition Eyes: Independence-Interdependence | -0.06 | 0.12 | -0.53 | .59 |
| Condition Camera: Independence-Interdependence | -0.29 | 0.14 | -2.09 | .03* |

Table A8. Aggregate Analysis. Poisson GLM Examining Destruction rates in the JoD by Condition, Independence-Inter-dependent Self-Construal

| | *B* | *SE B* | *Z* | *P* |
|---|---|---|---|---|
| Constant | 0.97 | 0.04 | 22.98 | <.001 *** |
| Condition: Eyes | -0.07 | 0.06 | -1.10 | .27 |
| Condition: Camera | -0.26 | 0.07 | -3.74 | <.001*** |
| Independence-Interdependence | 0.08 | 0.04 | 1.72 | .09 . |
| Condition Eyes: Independence-Interdependence | 0.09 | 0.07 | 1.33 | .18 |
| Condition Camera: Independence-Interdependence | -0.18 | 0.07 | -2.40 | .02 * |

*Figure A5. Study 1. The influence of Independent-Interdependent Self-Construal on Destruction Rates in the Joy of Destruction Game.*

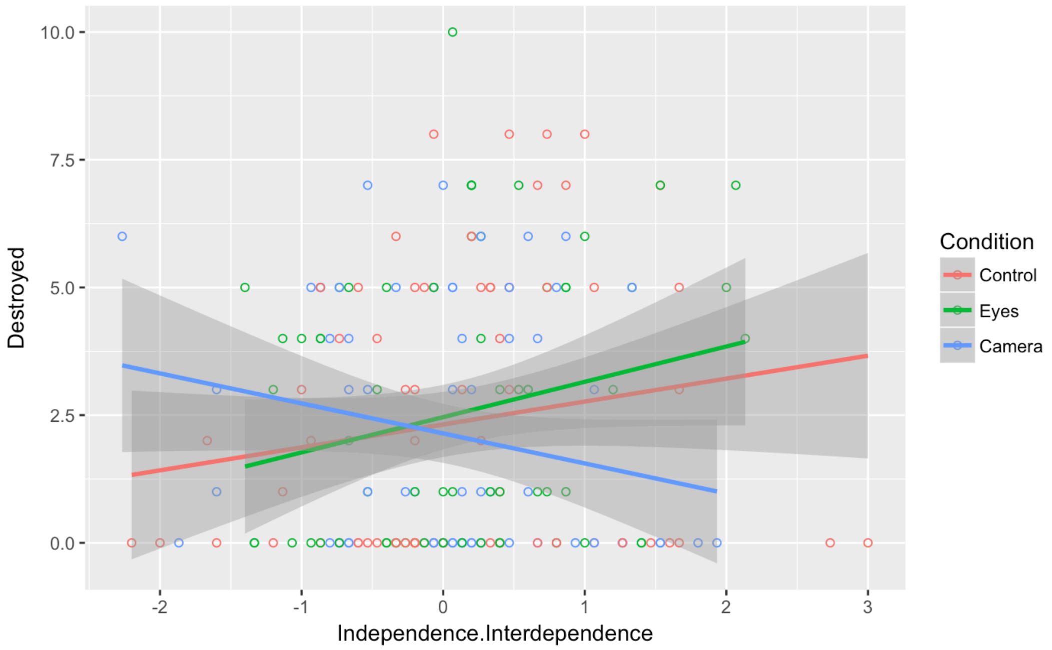


*Figure A6. Study 2. The influence of Independent-Interdependent Self-Construal on Destruction Rates in the Joy of Destruction Game.*

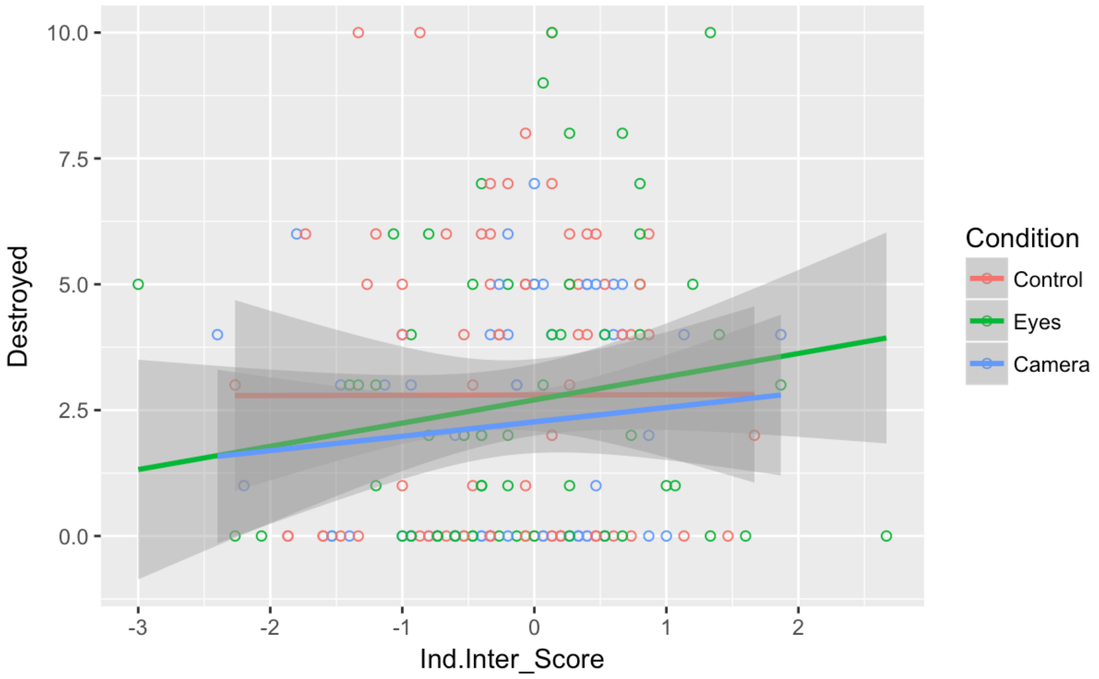


*Figure A7. Study 3. The influence of Independent-Interdependent Self-Construal on Destruction Rates in the Joy of Destruction Game.*

# 9. Appendices

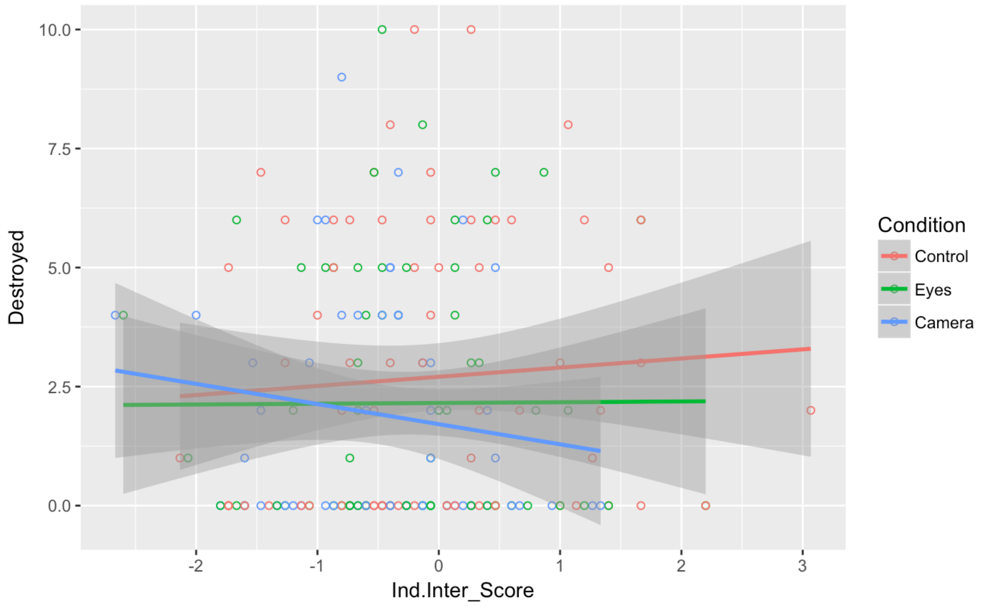
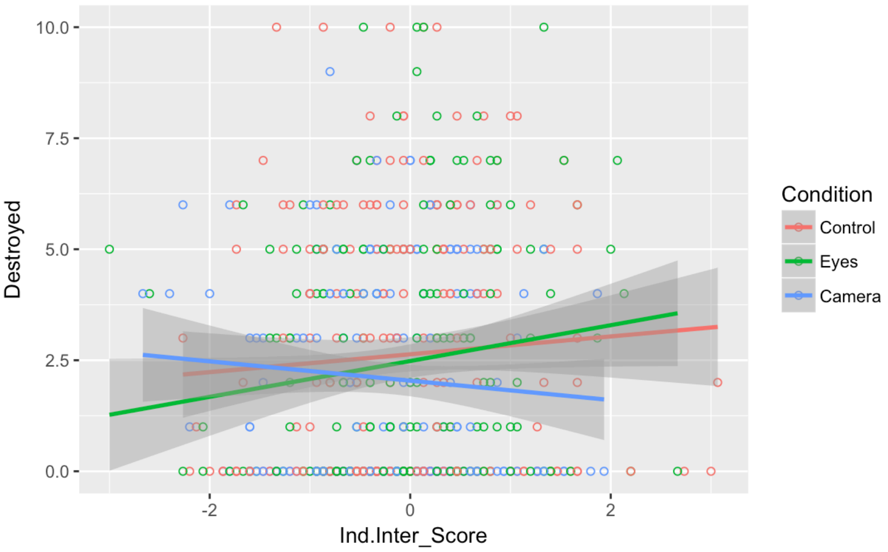


*Figure A8. Aggregate Analysis. The influence of Independent-Interdependent Self-Construal on Destruction Rates in the Joy of Destruction Game.*



### *Social Value Orientation Scores Analysis for Studies 1-3 & all 3 Studies in the Aggregate.*

Table A9. Study 1. Poisson GLM Examining Destruction rates in the JoD by Condition, SVO Score

| | *B* | *SE B* | *Z* | *P* |
|---|---|---|---|---|
| Constant | 1.82 | 0.49 | 3.71 | <0.001*** |
| Condition: Eyes | 1.76 | 0.67 | 2.62 | 0.009** |
| Condition: Camera | -0.94 | 0.67 | -1.40 | 0.16 |
| SVO Score | -0.02 | 0.01 | -1.96 | 0.05* |
| Condition Eyes: SVO Score | -0.04 | 0.02 | -2.48 | 0.01* |
| Condition Camera: SVO Score | 0.02 | 0.02 | 1.22 | 0.22 |

Table A10. Study 2. Poisson GLM Examining Destruction rates in the JoD by Condition, SVO Score

| | *B* | *SE B* | *Z* | *P* |
|---|---|---|---|---|
| Constant | 2.34 | 0.36 | 6.51 | <0.001*** |
| Condition: Eyes | -1.31 | 0.57 | -2.28 | 0.02* |
| Condition: Camera | -2.67 | 0.86 | -3.11 | 0.002** |
| SVO Score | -0.03 | 0.01 | -3.64 | <0.001*** |
| Condition Eyes: SVO Score | 0.03 | 0.01 | 2.26 | 0.02* |
| Condition Camera: SVO Score | 0.06 | 0.02 | 2.92 | <0.01** |

Table A11. Study 3. Poisson GLM Examining Destruction rates in the JoD by Condition, SVO Score

| | *B* | *SE B* | *Z* | *P* |
|---|---|---|---|---|
| Constant | 0.91 | 0.16 | 5.85 | <0.001*** |
| Condition: Eyes | -0.27 | 0.21 | -1.28 | 0.20 |
| Condition: Camera | -0.18 | 0.23 | -0.78 | 0.44 |
| SVO Score | 0.001 | 0.003 | 0.31 | 0.58 |
| Condition Eyes: SVO Score | 0.001 | 0.003 | -0.82 | 0.76 |
| Condition Camera: SVO Score | -0.002 | 0.004 | -0.82 | 0.41 |

Table A12.
Aggregate Analysis. Poisson GLM Examining Destruction rates in the JoD by Condition, SVO Score

| | *B* | *SE B* | *Z* | *P* |
|---|---|---|---|---|
| Constant | 1.01 | 0.10 | 9.88 | <0.001*** |
| Condition: Eyes | -0.13 | 0.14 | -0.95 | 0.34 |
| Condition: Camera | -0.20 | 0.16 | -1.29 | 0.20 |
| SVO Score | -0.001 | 0.002 | -0.48 | 0.64 |
| Condition Eyes: SVO Score | 0.001 | 0.003 | 0.52 | 0.60 |
| Condition Camera: SVO Score | -0.001 | 0.003 | -0.21 | 0.83 |

*Figure A9. Study 1. The influence of Social Value Orientation on Destruction Rates in the Joy of Destruction Game.*

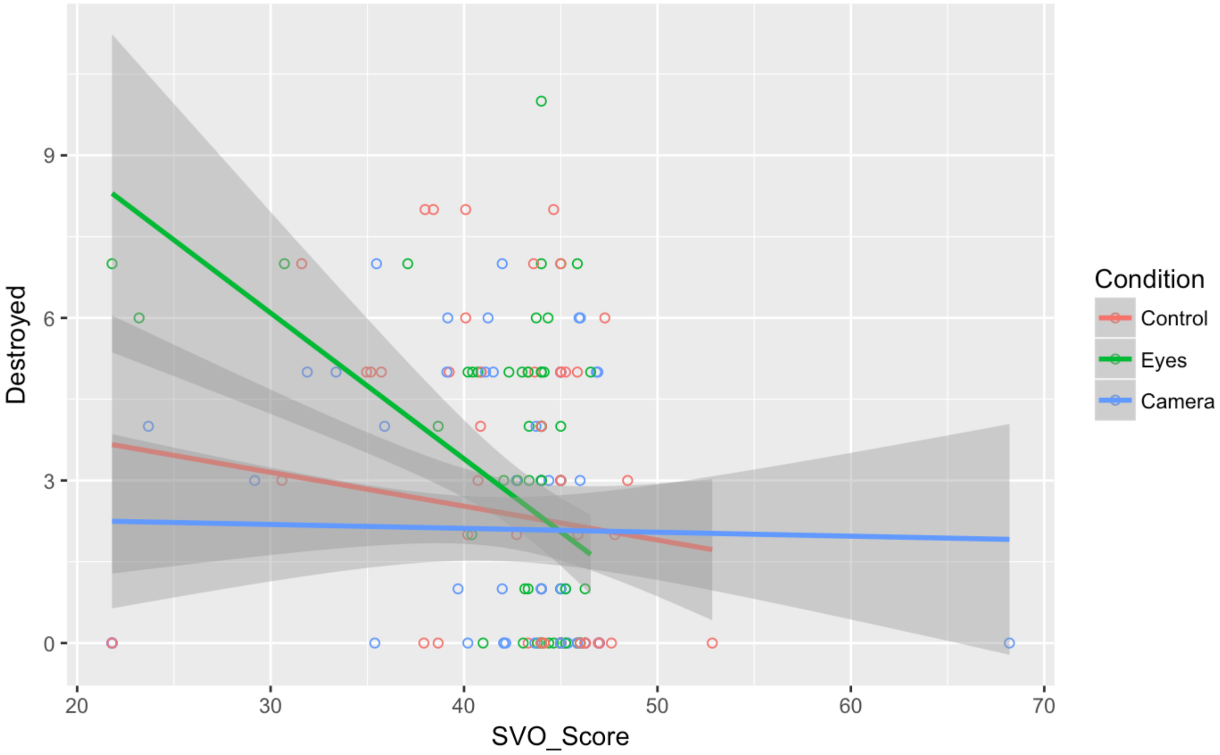


*Figure A10. Study 2. The influence of Social Value Orientation on Destruction Rates in the Joy of Destruction Game.*

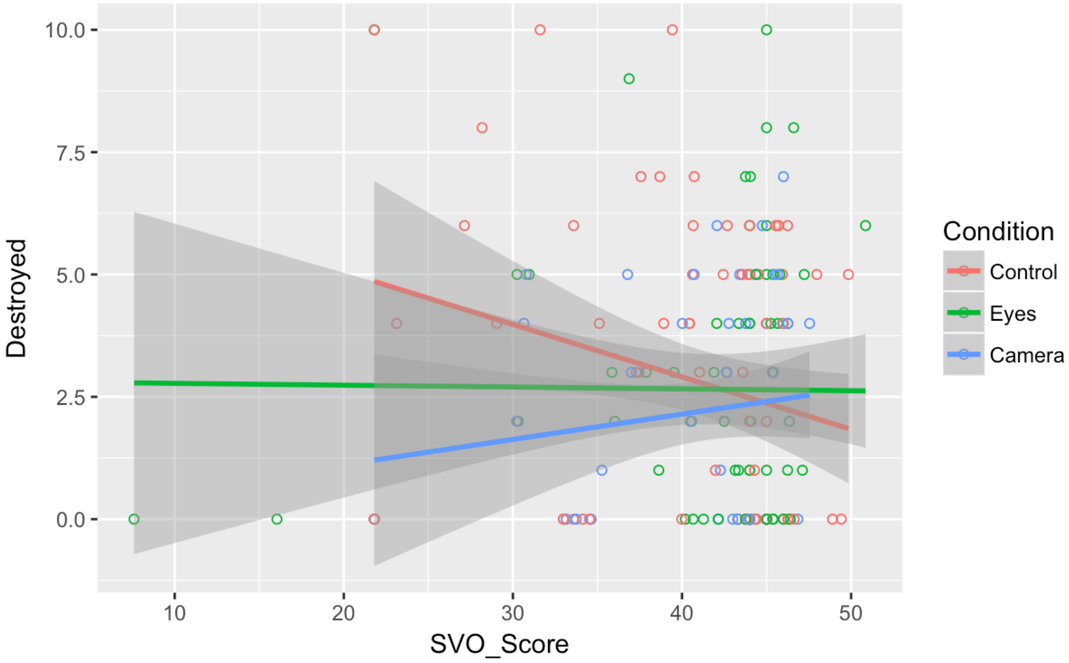


*Figure A11. Study 3. The influence of Social Value Orientation on Destruction Rates in the Joy of Destruction Game.*

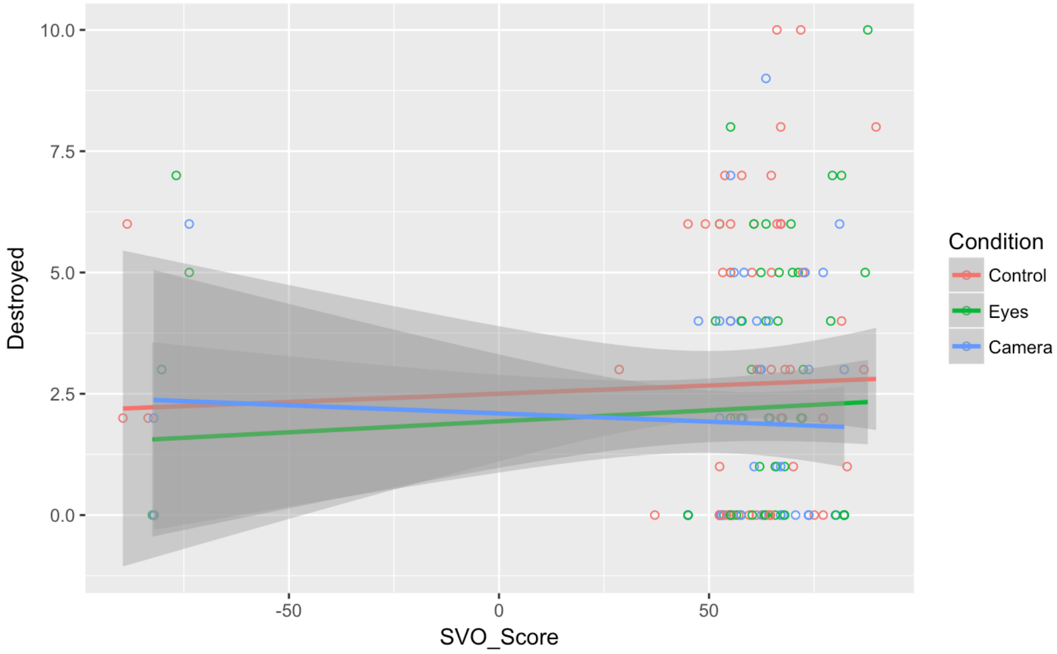
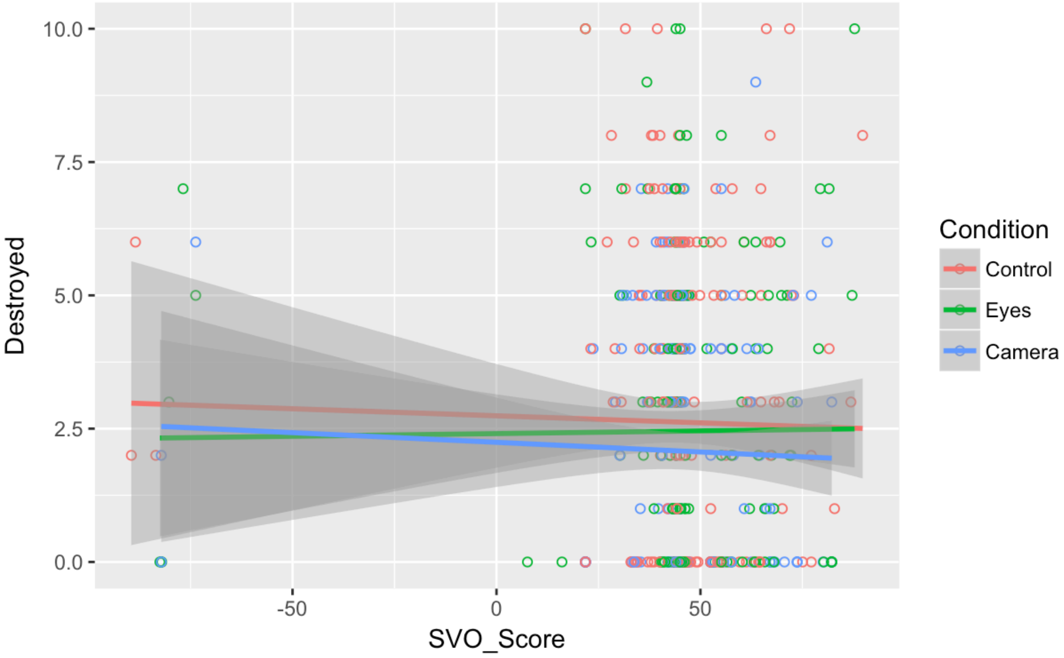


*Figure A12. Aggregate Analysis. The influence of Social Value Orientation on Destruction Rates in the Joy of Destruction Game.*



## *Centrality of Religion Scores Analysis for Studies 1-3 & all 3 Studies in the Aggregate.*

Table A13. Study 1. Poisson GLM Examining Destruction rates in the JoD by Condition, Religiosity

| | *B* | *SE B* | *Z* | *P* |
|---|---|---|---|---|
| Constant | 1.16 | 0.23 | 5.08 | <0.001*** |
| Condition: Eyes | 0.67 | 0.33 | 2.03 | 0.04* |
| Condition: Camera | -0.81 | 0.33 | -2.47 | 0.01* |
| Centrality of Religion Score (CRS) | -0.15 | 0.11 | -1.37 | 0.17 |
| Condition Eyes: CRS | -0.35 | 0.17 | -2.02 | 0.04* |
| Condition Camera: CRS | 0.34 | 0.15 | 2.26 | 0.02* |

Table A14. Study 2. Poisson GLM Examining Destruction rates in the JoD by Condition, Religiosity

| | *B* | *SE B* | *Z* | *P* |
|---|---|---|---|---|
| Constant | 1.16 | 0.17 | 6.68 | <0.001*** |
| Condition: Eyes | -0.46 | 0.25 | -1.82 | 0.07 . |
| Condition: Camera | -0.72 | 0.30 | -2.41 | 0.41 |
| Centrality of Religion Score (CRS) | -0.06 | 0.07 | -0.82 | 0.41 |
| Condition Eyes: CRS | 0.18 | 0.10 | 1.76 | 0.08 . |
| Condition Camera: CRS | 0.23 | 0.13 | 1.80 | 0.07 . |

Table A15. Study 3. Poisson GLM Examining Destruction rates in the JoD by Condition, Religiosity

| | *B* | *SE B* | *Z* | *P* |
|---|---|---|---|---|
| Constant | 0.49 | 0.18 | 2.69 | <0.01** |
| Condition: Eyes | 0.21 | 0..28 | 0.73 | 0.47 |
| Condition: Camera | -0.51 | 0.30 | -1.67 | 0.09 . |
| Centrality of Religion Score (CRS) | 0.23 | 0.07 | 3.09 | 0.002** |
| Condition Eyes: CRS | -0.20 | 0.12 | -1.58 | 0.11 |
| Condition Camera: CRS | 0.04 | 0.11 | 0.38 | 0.71 |

Table A16. Aggregate Analysis. Poisson GLM Examining Destruction rates in the JoD by Condition, Religiosity

| | *B* | *SE B* | *Z* | *P* |
|---|---|---|---|---|
| Constant | 0.89 | 0.11 | 8.04 | <0.001*** |
| Condition: Eyes | 0.01 | 0.16 | 0.08 | 0.93 |
| Condition: Camera | -0.60 | 0.17 | -3.45 | 0.0006*** |
| Centrality of Religion Score (CRS) | 0.03 | 0.05 | 0.73 | 0.47 |
| Condition Eyes: CRS | -0.04 | 0.07 | -0.53 | 0.60 |
| Condition Camera: CRS | 0.17 | 0.07 | 2.32 | 0.02* |

*Figure A13. Study 1. The influence of Centrality of Religion on Destruction Rates in the Joy of Destruction Game.*

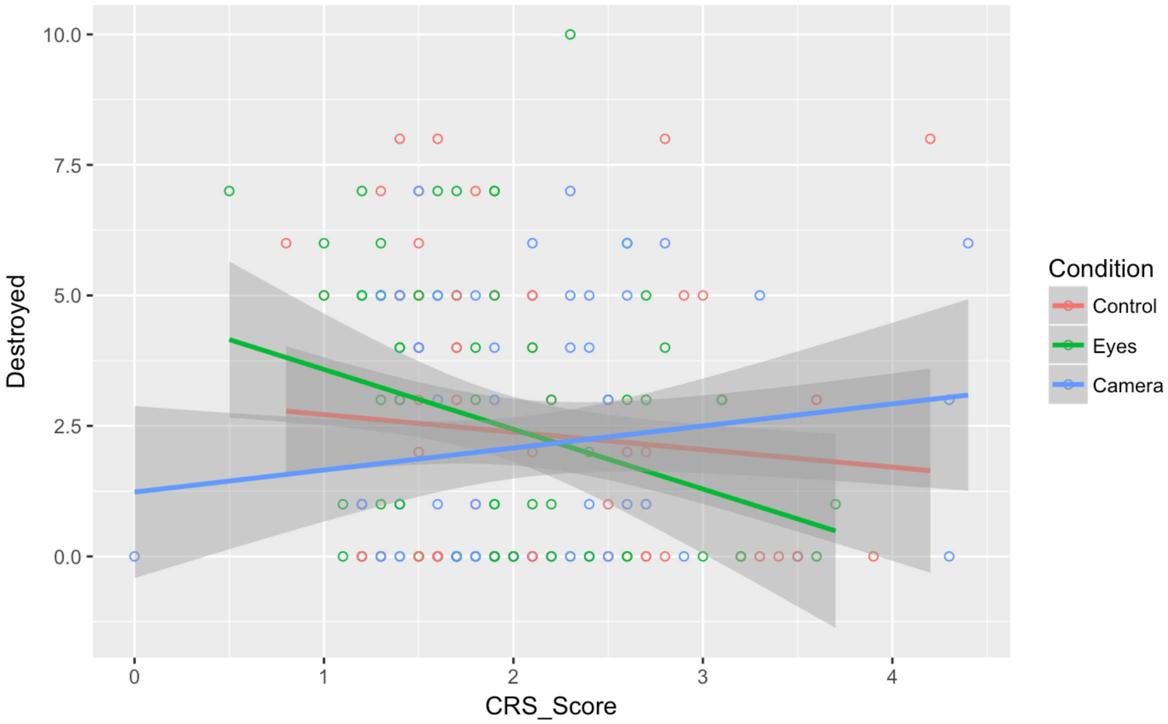


*Figure A14. Study 2. The influence of Centrality of Religion on Destruction Rates in the Joy of Destruction Game.*

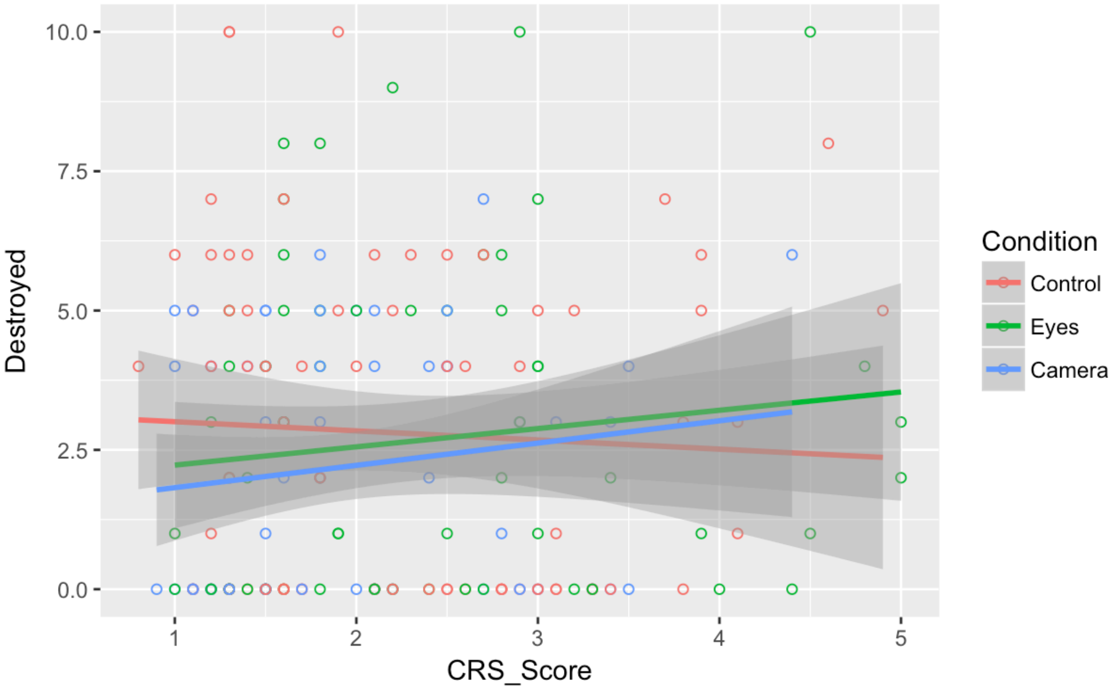


*Figure A15. Study 3. The influence of Centrality of Religion on Destruction Rates in the Joy of Destruction Game.*

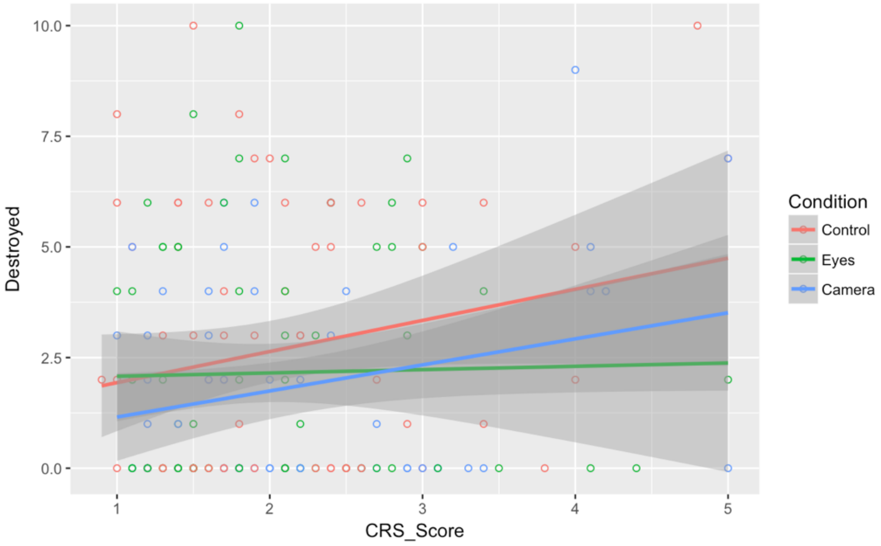
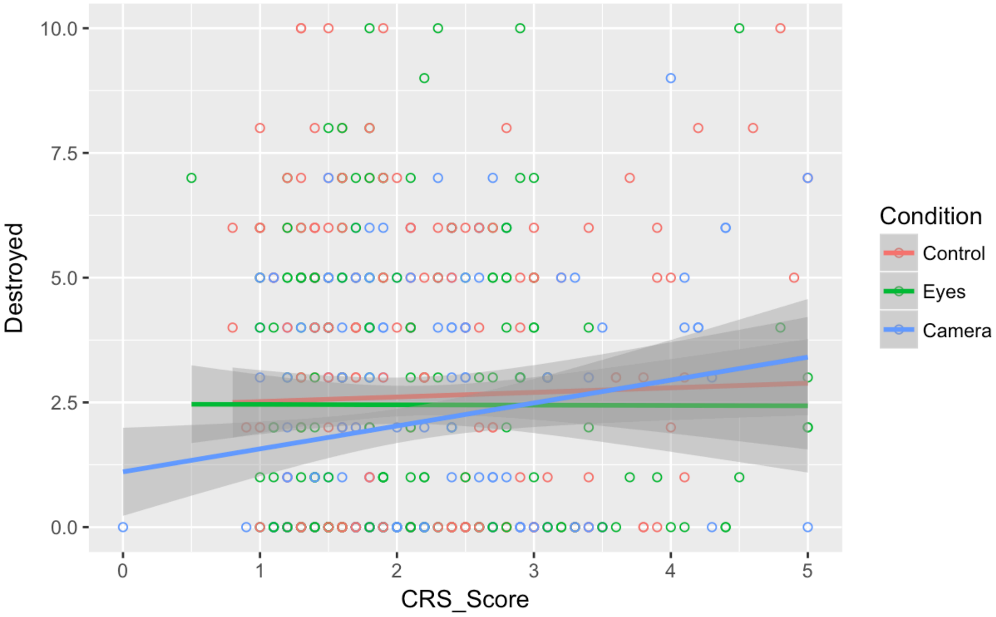


*Figure A16. Aggregate Analysis. The influence of Centrality of Religion on Destruction Rates in the Joy of Destruction Game.*



# C

# Appendix C for Chapter 7

Table C.1. (Overleaf) Provides the summary statistics for all the studies included in the final meta-analysis in Chapter 7.

Table C1. Final Meta-Analysis

| | | | Eye Cues Present | | | Control | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Citation & Experiment Number** | **Outcome Variable** | ***N*** | **ASB Yes** | **ASB No** | ***N*** | **ASB Yes** | **ASB No** | $OR_{log}$ | $Var_{LOR}$ | $SE_{LOR}$ | **95% CI Lower** | **95% CI Upper** | ***Z*** | ***P*** |
| **1** | Ernest-Jones et al.(2011) | Litter | 273 | 56 | 233 | 289 | 98 | 175 | -0.85 | 0.04 | 0.20 | -1.23 | -0.46 | -4.34 | **0.01** |
| **2** | Bateson et al.(2013) | Litter | 305 | 80 | 225 | 297 | 83 | 214 | -0.09 | 0.03 | 0.18 | -0.45 | 0.27 | -0.47 | 1.28 |
| **3** | Baillon et al.(2013), Experiment 1 | Destruction of others money | 51 | 9 | 42 | 49 | 19 | 30 | -1.08 | 0.22 | 0.47 | -2.00 | -0.16 | -2.31 | 0.57 |
| **4** | Zengerink (2013), Experiment 2 | Litter | 630 | 297 | 333 | 314 | 170 | 144 | -0.28 | 0.02 | 0.14 | -0.55 | -0.01 | -2.02 | 0.78 |
| **5** | Bateson et al.(2015), Experiment 1 | Litter | 147 | 8 | 139 | 137 | 23 | 114 | -1.25 | 0.18 | 0.43 | -2.10 | -0.41 | -2.92 | 0.23 |
| **6** | Bateson et al.(2015), Experiment 2 | Litter | 216 | 31 | 185 | 97 | 25 | 72 | -0.73 | 0.09 | 0.30 | -1.32 | -0.14 | -2.41 | 0.50 |
| **7** | Cai et al.(2015), Experiment 1 | Dishonesty | 66 | 30 | 36 | 65 | 31 | 34 | -0.09 | 0.12 | 0.35 | -0.78 | 0.60 | -0.26 | 1.17 |
| **8** | Cai et al.(2015), Experiment 3 | Dishonesty | 64 | 18 | 46 | 66 | 22 | 44 | -0.25 | 0.15 | 0.38 | -0.99 | 0.50 | -0.64 | 1.34 |
| **9** | Hoffman et al.(2015) | Dishonesty | 20 | 7 | 13 | 20 | 9 | 11 | -0.42 | 0.42 | 0.65 | -1.69 | 0.85 | -0.64 | 1.34 |
| **10** | Palomäki et al.(2015) | Dishonesty | 105 | 13 | 92 | 86 | 9 | 77 | 0.19 | 0.21 | 0.46 | -0.71 | 1.09 | 0.41 | 0.69 |
| **11** | Huang et al.(2015), Experiment 2 | Corruption | 24 | 10 | 14 | 24 | 9 | 15 | 0.17 | 0.35 | 0.59 | -0.98 | 1.33 | 0.30 | 0.78 |
| **12** | Huang et al.(2015), Experiment 3 | Corruption | 30 | 11 | 19 | 31 | 24 | 7 | -1.78 | 0.33 | 0.57 | -2.90 | -0.66 | -3.11 | 0.17 |
| **13** | Meleady et al. (2017), Experiment 1 | Engine idling | 112 | 83 | 21 | 104 | 82 | 39 | 0.63 | 0.10 | 0.31 | 0.02 | 1.24 | 2.02 | **0.04** |
| **14** | Dear, Dutton, Fox, Joy of Destruction (1) | Destruction of others money | 70 | 176 | 524 | 71 | 168 | 542 | 71 | 0.08 | 0.02 | 0.12 | -0.16 | 0.32 | 0.65 |
| **15** | Dear, Dutton, Fox, Joy of Destruction (2) | Destruction of others money | 71 | 191 | 519 | 77 | 215 | 555 | 77 | -0.05 | 0.01 | 0.12 | -0.28 | 0.18 | -0.44 |
| **16** | Dear, Dutton, Fox, Joy of Destruction (3) | Destruction of others money | 66 | 142 | 518 | 70 | 188 | 512 | 70 | -0.29 | 0.02 | 0.13 | -0.54 | -0.04 | -2.29 |

| | *Continued…* | | **Experimental Locations** | | | **Control Locations** | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Citation & Experiment Number** | **Outcome Variable** | ***N*** | **Frequency of ASB Before** | **Frequency of ASB After** | ***N*** | **Frequency of ASB Before** | **Frequency of ASB After** | $OR_{log}$ | $Var_{LOR}$ | $SE_{LOR}$ | **95% CI Lower** | **95% CI Upper** | ***Z*** | ***P*** |
| **17** | Nettle et al.(2012) | Bicycle Theft | N/A | 31 | 51 | N/A | 39 | 15 | -1.45 | 0.14 | 0.38 | -2.20 | -0.71 | -3.83 | **0.04** |
| **18** | Keep Britain Tidy (2014) | Dog Fouling | N/A | 861 | 434 | N/A | 2159 | 1208 | 0.10 | 0.00 | 0.07 | -0.03 | 0.24 | 1.51 | 0.13 |
| **19** | Dear, Dutton, Fox, Ealing Cycle Crime | Cycle Crime | N/A | 30 | 15 | N/A | 30 | 11 | -0.31 | 0.22 | 0.47 | -1.24 | 0.62 | -0.65 | 1.34 |
| **20** | Dear, Dutton, Fox, Hereford Eyes on Crime | Crime & ASB | N/A | 310 | 460 | N/A | 385 | 385 | -0.39 | 0.01 | 0.10 | -0.60 | -0.19 | -3.83 | **0.03** |

# D

# Appendix D for Chapters 1-7

All supporting materials to Chapters 1-7 can be found via the Centre for Open Science's online repository: www.osf.io

The Centre for Open Science is non-profit open source software project that facilitates open collaboration in science research. The framework is being employed widely in seeking to improve the reproducibility of psychological research[3].

Data includes all data files, R Code and .pdf R Output for all studies reported herein. The data is stored in the project folder entitled 'Appendix D to KPD Thesis 'Towards a Psychology of Surveillance''.

The link in full is: https://osf.io/rhzgt/?view_only=c31a90486177469c935806609aba2612

---

[3] See: Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349*(6251), aac4716. doi: 10.1126/science.aac4716 & Yong, E. (2012). Bad copy. *Nature*, *485*(7398), 298-300. doi:10.1038/485298a