Understanding and Addressing Cultural Variation in Costly Antisocial Punishment

Dr. Joanna J. Bryson

University of Bath
Department of Computer Science
Claverton Down
Bath, United Kingdom BA2 7AY

Report Date: August 2012

Final Report for 01 July 2010 to 31 December 2011

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From – To)*
15 August 2012	Final Report	1 July 2010 – 31 December 2011

4. TITLE AND SUBTITLE

Understanding and Addressing Cultural Variation in Costly Antisocial Punishment

5a. CONTRACT NUMBER

FA8655-10-1-3050

5b. GRANT NUMBER

Grant 10-3050

5c. PROGRAM ELEMENT NUMBER

61102F

6. AUTHOR(S)

Dr. Joanna J. Bryson

5d. PROJECT NUMBER

5d. TASK NUMBER

5e. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Bath
Department of Computer Science
Claverton Down
Bath, United Kingdom BA2 7AY

8. PERFORMING ORGANIZATION REPORT NUMBER

N/A

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

EOARD
Unit 4515 BOX 14
APO AE 09421

10. SPONSOR/MONITOR'S ACRONYM(S)

AFRL/AFOSR/RSW (EOARD)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

AFRL-AFOSR-UK-TR-2012-0023

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited. (approval given by local Public Affairs Office)

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This report results from a contract tasking University of Bath as follows: The varieties of human cultures are a rich treasure, valued for their wealth of knowledge as well as their aesthetics. However, some aspects of cultural variation result in measurable metric differences in indices with apparently universal ethical value, for example infant mortality or literacy rates. We seek here to understand the cultural traits that lead members of some societies to be more likely than others to employ strategies of economically-advantageous altruistic collaboration. We focus on explaining the documented propensity in some cultures (e.g. in Greece, Russia and the Middle East) for antisocial punishment — punishing those who are significantly more altruistic than the punisher, even though the punisher benefits from their altruism. The results of this research project indicate that anti-social punishment may indeed occur in contexts where other participants are not mutually-acknowledged members of trusted group, yet generosity may in absence of other information be taken as an indication that in fact trust is merited. We have found that costly punishment is best understood as having impact not only on global economics but also on individual competition, and that the apparently maladaptive behaviour of anti-socially punishing those more generous than ourselves may even in some contexts be a sensible response, when our own well-being is (or at least appears to be) most determined by our relative dominance to our local neighbours, not to how well our neighbourhood performs as a whole.

15. SUBJECT TERMS

EOARD, Behavioral Science, Psychology

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18, NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT UNCLAS	b. ABSTRACT UNCLAS	c. THIS PAGE UNCLAS	SAR	35	JAMES H. LAWTON, Ph.D.
					19b. TELEPHONE NUMBER *(Include area code)* +44 (0)1895 616459

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39-18

Final Report

Title: **Understanding and Addressing Cultural Variation in Costly Antisocial Punishment**
Grant Number: FA8655-10-1-3050

Joanna J. Bryson[†], PI, with Benedikt Herrmann[‡], CI

† Dept. of Computer Science, University of Bath, BA2 7AY, United Kingdom
‡ School of Economics, University of Nottingham, NG7 2RD, United Kingdom

Directorate:	Mathematics, Information and Life Sciences
Program:	Collective Behavior and Socio-Cultural Modeling
Program Manager:	Dr. Joseph B. Lyons (originally Dr. Terence Lyons)
Local PM:	Dr. James Lawton (originally Lt. Col. Tammy Savoie, intermediately Col. Lisa Boyce)
Report Period	1 July 2010 – 31 December 2011 (incl. 6 month NCE)

Abstract

This effort was designed as the first year of a three-year project with two concurrent research goals:

1. to improve the scientific standard of application of agent-based modelling (ABM) to the social sciences, and

2. to explain cultural variation by global region in the economically-inexplicable propensity of some actors to punish those who contribute more to the public good than the actor (anti-social punishment, ASP).

The aims of this first year were to build a team, build research tools, review extant literature and findings, establish hypotheses for testing and achieve HSP. Primary scientific findings:

1. All populations show some individuals exhibiting ASP. Variation is due to difference in proportions of actor types, not global shifts of behaviour per region.

2. ASP appears to be more likely when the generous benefactor victim is perceived as being "out group" by the actor applying ASP.

3. ASP in itself is an unsustainable strategy, disadvantageous for an individual unless it correlates with some other compensating advantage, probably social dominance.

4. ASP negates currently-dominant theories for explaining the extent of human cooperation.

We also developed a reproducible research software framework incorporating ABM, and achieved HSP from our IRB three times, but never from the USAF SG.

Final Report: Understanding and Addressing Cultural Variation in Costly Antisocial Punishment

Joanna J. Bryson

1 Introduction

The variety of human cultures is one of the joys of contemporary human life. However, our respect and appreciation for diversity does not stop us from observing that in some aspects, cultural variation can result in measurable metric variation on indices, some of which have nearly-universal cross-cultural appeal, for example reducing infant mortality or increasing literacy. The primary goal of this research project was to study cultural variation in one such trait: the propensity of group members to discover mechanisms to optimise altruistic economic collaboration. This variation was discovered and published in *Science* by one of this project's investigator's and his colleagues (Herrmann et al., 2008a). Although the data is based on formal laboratory experiments where participants play a 'game' for money, the results correlate highly with national GDP, leading us to the possibility that the behaviour measured in the laboratory may have fundamental impact on the economic well-being of a nation. Further, the variation between cultures is not arbitrary, but rather seems to be clustered by global region. Thus Boston, several cities in Northern Europe, the Far East, and Melbourne show high levels of profitable and altruistic economic collaboration, while Athens, Istanbul, regions of the Middle East and of the former Soviet Republics show relatively low levels.

This project also had a secondary, methodological goal pursued simultaneously with the social science goal. This goal—to improve the scientific application of agent-based modelling (ABM)—was actually originally the driving motivation for the Programme Manager Dr. Terence Lyons soliciting the present proposal from Dr. Bryson. Given the high level of AFOSR personell turn over during the period of this grant, and also to clarify the narrative and scientific contributions of the research, I will begin with a brief overview of the timeline for this project. I will then review the scientific context in which the primary research was done, and the primary scientific outcomes. I will then summarise policy implications of these outcomes. Next, I will review the outcomes relevant to the secondary goal concerning ABM. Finally, I include a list of dissemination (talks and articles) for the project outcomes to date. For the convenience of the reader, these sections are hyperlinked in the following Table of Contents.

Contents

2 Brief Project Timeline

This is a timeline of the administrative highlights of this project, including hiring and stages of the HSP process. It is based primarily on email records.

- **26–28 March 2008**: Terence Lyons approaches Joanna Bryson at the *AAAI Spring Symposium on Emotion, Personality, and Social Behavior*, held at Stanford, where Bryson is presenting work on synthesis of realistic human-like emotions. Lyons describes the problem of the scientific application of ABM to the social sciences in the context research he has previously funded. This topic is not relevant to the symposium, but is relevant to Bryson et al. (2007)'s work on the scientific application of ABM to biology. Over subsequent email conversation, Lyons suggests Bryson submits a whitepaper for a BAA in socio-cultural modelling, but that she must find an 'anthropologist' to work with.

- **11 April 2008** Bryson first emails Dr. Benedikt Herrmann with the proposal to collaborate, in response to his article in *Science*. She subsequently meets Herrmann and the senior academic on the work, Simon Gächter.

- **14 May 2008** Bryson and Herrmann submit a whitepaper in consultation with Gächter, *Understanding and Addressing Cultural Factors Leading to Costly Antisocial Punishment*.

- **7 June 2008** Lyons solicits a full proposal, saying "the study of the evolution of altruistic collaboration is timely."

- **30 July 2008** AFOSR acknowledges receipt of proposal. In personal communication, Lyons says that ordinarily results would be known by October, but that in this case with the coming election they may be delayed until December.

- **August 2008** Bryson is approached by Harvey Whitehouse of Oxford to assist with ABM on his *Explaining Religion* project as a consultant. She is formally a visiting research fellow to Oxford Anthropology for the entire year of 2009. She runs a seminar series for his group bringing in people with top ABM publications on cultural evolution to discuss the role of religion in cultural change. Daniel Taylor, then a Bath MSc student, also attends these meetings.

- **September 2008** Herrmann turns down positions at Oxford and Nottingham to accept one at the European Commission, in a DG focussed on fostering research in Europe. He maintains his fellow position at Nottingham and continues writing and directing research in his available time.

- **18 February 2009** Lyons sends Bryson review comments and asks for a reply, which is provided. By telephone he says that the reviews were very strong but the cost is high and should be cut. In the time since the proposal was submitted, the pound has dropped against the dollar by 40%, but UK academic salaries have increased about 7%. Lyons also mentions that HSP might be a problem; Bryson assures him that she knows the procedure at Bath and has never had a problem with it. Unfortunately, this is only for Bath approval — no one currently in Bath has experienced contemporary US Military HSP procedures or knows what they are up against.

- **21 April 2009** Lyons acknowledges receipt of the new version of the proposal. Although there is no formal notice of contract, he says in email that no further review is needed and funds can be transferred as soon as HSP is achieved. Bryson is already trying to find the UK regional IRB Bath is designated to use (it has changed since the last time Bath needed US HSP).. The IRB proves to be biomedically oriented and to require extensive documentation beyond the AFOSR's guidelines.

- **1 September 2009** HSP request is filed with the local IRB in advance of their first meeting after their summer break.

- **5 October 2009** The IRB declares the proposed research does not require HSP.

- **8 October 2009** The IRB agrees to reconsider the case in their 23 October meeting, which Bryson attends to answer any questions.

- **17 November 2009** Bryson receives an updated permissions letter from the IRB. The original letter documenting the award of *exempt* status had been sent the previous week, but contained a misphrased condition, which after consultation with AFOSR HSP contractor Stephanie Bruce, Bryson asked to be clarified. The completed application for HSP is sent surface-mail to Bruce.

- **7 January 2010** Bruce emails Bryson requesting the documents that have already been surface mailed. Bryson, working from abroad during a break in term, sends Bruce all the documents she has on her laptop. Apparently due to issues with large attachments, it takes until 29 January for Bruce to successfully receive and download these documents.

University of Bath, Department of Computer Science *J J Bryson, PI*

- **12 January 2010** London Liaison Officer Tammy Savoie formally acknowledges receipt of the full, three year, $927,315 proposal, then asks for a second proposal to be filled in for the first 12 months.

- **4 March 2010** Savoie notifies Bryson that funding has been secured and that a check will be sent.

- **19 March 2010** Bruce sends Bryson and Savoie a list of additional HSP requirements from the SGEC. Bryson and Savoie ask (repeatedly, over weeks and months) for clarification. Clarification is ultimately provided in November in a teleconference with Bruce and Malcolm Barth of the SG.

- **8 April 2010** Savoie suggests (in light of ongoing lack of clarification from Bruce) the removal of the human subject research from the first 12 months contract and making HSP the task of the postdoctoral researcher so that the funding can advance.

- **June 2010** Contract is signed, positions are advertised.

- **1 July 2010** Formal start of grant. Lyons had recommended writing the grant to start research 3 months *after* formal start to allow time for hiring take place.

- **30 July 2010** The University of Bath awards Daniel Taylor a PhD studentship (full three years of tuition plus a three-year stipend) based on award of contract and as a sign of support for the research.

- **10 August 2010** *The Boston Globe* breaks the story of Marc Hauser's academic misconduct involving forging scientific results. Bryson, a former Hauser postdoc, resolves to include anti-fraud measures in the general methodological aspect of the grant.

- **4 October 2010** Grant kickoff day with newly started postdocs Karolina Sylwester and Pablo Lucas. Also attending are programmer James Mitchell, recently hired but not starting until November, Investigators Bryson and Herrmann, PhD students Taylor, Marios Richards and Gideon Gluckman and visiting research fellow Dr. Will Lowe. Richards, Gluckman and Lowe are not funded through the grant, but Lowe and Gluckman ultimately make contributions to it.

- **8–13 November** Ken Boff visits Bath in Lyons' place, meets the team. Lyons is too ill to travel. Savoie visits Bath, she, Bryson and Sylwester have teleconference with Bruce and Malcolm Barth (SG), get a new list of SG requirements (actually, Barth forwards the original list he had sent Bruce at the beginning of the year, but they are new to Bath). Savoie, Boff and Bryson then attend an AFOSR-funded meeting in Plymouth, *Interdisciplinary Workshop on Society, Culture and Language*.

- **December 2010** Sylwester submits to the IRB a "Request for substantial amendment" to Bath's existing HSP status. The amendments are required both for meeting the SG's requirements and because research has not started over 12 months after the original assessment was granted, and have required significant work with Bath University lawyers and administration to meet SG requirements. The IRB request modification of one of the documents, this is sent 7 January 2011.

- **11 January 2011** Bruce first mentions an on-line HSP training course and suggests Sylwester should take it. Sylwester requests the URL but never receives it.

- **16 January 2011** Savoie notifies Bryson that Lisa Boyce is taking over Savoie's role in London with respect to this project.

University of Bath, Department of Computer Science *J J Bryson, PI*

- **26 January 2011** Bath receive *exempt* status again from IRB, and check with Bruce that the letter is correctly phrased.

- **March 2011** Pablo Lucas leaves the project to take a postdoc near his partner in Helsinki, Finland.

- **2 May 2011** Simon Powers joins the project.

- **12 May 2011** Lisa Boyce informs Bryson by phone that funding will not be renewed. Following clarification, notice is given to all employees of the project. Powers is entitled to one month's notice, Sylwester and Mitchell to three.

- **20 May 2011** Joseph Lyons notifies Bryson of Terence Lyons' passing and that he is taking Lyons' role.

- **8 June 2011** Boyce notifies Bryson that SG disagrees with IRB conclusion that work is *exempt*. Boyce and Bryson query why Bryson was never asked to take the on-online HSP training course. Bryson asks the IRB to update the status and forms the next day.

- **21 June 2011** Boyce notifies Bryson of acceptance of NCE, allowing staff to be paid for the period of their notice out of funds saved by late hire of Mitchell, non-payment of Herrmann (due to difficulties with EC), and months between Lucas' leaving and Powers' arrival. Department of Computer Science provides additional funding so that Powers is able to be paid for three months as well, and team is kept together until the end of September.

- **22 June 2011** IRB returns assessment of *exempt* again. Bryson immediately petitions the IRB to fill in forms as *minimal risk* anyway since anything that qualifies for exempt also qualifies for minimal risk.

- **16 November 2011** Bryson receives notification from IRB of classification of *minimal risk*, which is dated 12 September. Further delay was apparently caused by the need under minimal risk to obtain clearance to extract minutes for sending to the AFOSR. In her response to notice of this, Boyce informs Bryson that James Lawton has taken over her role with respect to this project. Bryson suggests to Lawton that the HSP process be dropped and to proceed directly to final report.

- **31 December 2011** Formal end of NCE.

3   Review of Scientific Background on Costly Punishment

Herrmann et al. (2008a) showed that in some human subject pools (e.g. university undergraduates in the Boston, Melbourne, Chengdu and Zurich) members tend to quickly exploit an experimental context in which mutual investment leads to mutual benefits. However, in other societies (e.g. university undergraduates in Muscat, Istanbul, Minsk and Athens) substantial proportions of the society will pay a penalty in order to further penalize others who are being more generous than themselves. This is despite the fact that this generosity is benefiting all group members other than the benefactor, including the punishers (See Figure 1). Such punishment of altruism is called *antisocial punishment* (ASP).

Herrmann et al. sought correlates for the prevalence of ASP in a culture, and found that the two strongest are that it inversely correlates to both Gross Domestic Product (GDP) and the Rule of Law (Kaufmann et al., 2004). Their article in *Science* suggests that "weak norms of civic cooperation and the weakness of the rule of law in a country are significant predictors of antisocial punishment. Our results show that punishment opportunities are socially beneficial only if complemented by

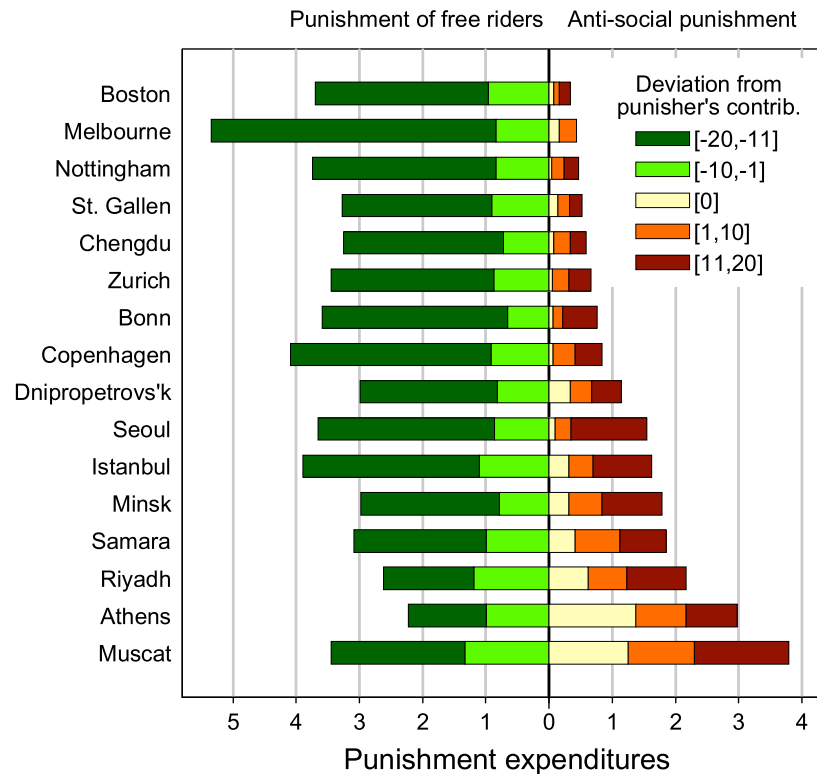Punishment of free riders   Anti-social punishment

Figure 1: Mean punishment expenditures for a given deviation from the punishers' contribution. Contributions are in "tokens", which are assigned equivalent monetary value across the cultures. After Herrmann et al. (2008a, figure 1, p. 1363).

strong social norms of cooperation." But correlation does not show causation. Can we be sure that the propensity for ASP does not itself lead to a weak rule of law? Or that both could be caused by some other factor? As will be discussed further below in Section 6, the theory vacuum caused by the disruptive findings of Herrmann et al. created a situation ripe for exploration with ABM.

3.1   The Data: How Altruism and Antisociality are Measured

All human subject data for this research was collected using a paradigm from the new branch of economics, generally called either *experimental economics* or *behavioural economics*. This is somewhat like classic experimental psychology, except a few special protocols are followed:

- The performance of all individuals must be rewarded by a sufficiently significant financial reward that there can be no doubt that financial motivation is present.

- There can be absolutely no deception of subjects. The standard psychological 'tricks' of telling subjects one thing is the task while measuring something else are not allowed. In fact, if there has been any deception committed against subjects in the subject pool (e.g. any history of non-payment or unexpected payment) the entire university is blacklisted.

- All subjects must demonstrate understanding of the complete consequences of their own and the other team member's actions. This is done by means of a test after training. Subjects that cannot pass the test are excluded from the experiment. Thus we know that subjects know exactly what they are doing, and what benefits or sacrifices they are making for themselves depending on their actions.

In cross-cultural experimental economics research, the players play for tokens, for which they know the value in local currency. The reason they play for tokens is that it is easier to reason about how to divide 20 tokens than (for example) $7.82. Thus the motivation is provided by local currency, but the reasoning is supported to be more similar across cultures.
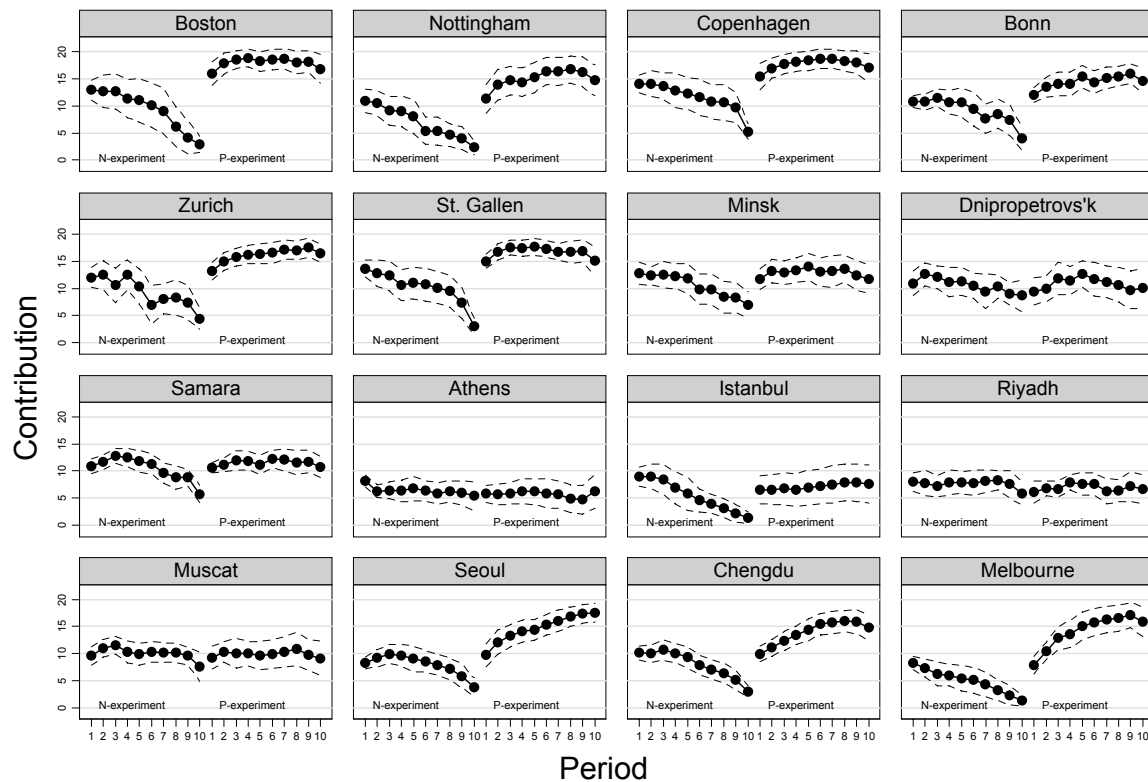
Figure 2: Development of average contributions in the **N**on-punishment and **P**unishment experiments, separately for each subject pool. The dashed lines indicate standard error of mean ($\pm SEM$). After Herrmann et al. (2008b, figure S4).

The standard type of behavioural economics experiment for assessing costly punishment is called the Public Goods Game (PGG). In the basic form of this game there is no punishment. PGG represents a social dilemma because individual interests are in conflict with the groups interests. In the standard form a group is determined by an experimenter, but members are not identified to each other and only interact by computer screens[1]. This anonymity is maintained to ensure group members do not act out of fear or expectation of retribution after the game. In a single round of PGG, each member is allocated 20 tokens, and then individuals are allowed to contribute any portion of their allocation to the public pool. Allocations to the public pool are multiplied by the

---

[1]In rural conditions the computers may be replaced with pen and paper for recording decisions, then the results are communicated to group members by the experimenter.

experimenter then divided equally between all group members. However, the multiplication is never so great that an individual receives as much money back from their own investment as they paid. Thus individuals who do not contribute anything or contribute less than others gain a financial advantage, at least for that round. PGG may be played as a single round, but for the results shown here they are played in ten rounds with the same group.

In the punishment condition, after a round of PGG individuals can anonymously punish others. This punishment is not based on identity, but only on the other's contribution to the most recent public pool. Importantly, subjects never learn any information about who punishes the, only about what others have contributed to the pool. In the studies shown here, the cost/effect ratio is that for every token a punisher pays, the punishee loses three tokens. When an individual punishes someone who has contributed less than they have, this punishment is termed 'altruistic' (AP) because the punisher pays a cost, yet the whole group benefits if (as seems often to be the case) this action leads cooperation (contributions to the public pool) to stabilise at a higher level. On the other hand, if punishers punish those who contribute more than they do, this is called 'anti-social'. Herrmann et al. (2008a) were the first to document societies with large amounts of ASP, and showed that this could in some cases completely counter the expected benefits of cooperation. In the Swiss contexts where these experiments were first run, the punishment condition of the PGG reliably resulted in a better economic outcome for the subjects, but this was not true in some societies with high levels of ASP.

In most of the data reported here (which is due to Herrmann et al.) subjects played two rounds of 10 PGG, one with punishment and one without. For most subject pools the order of the games (punishment or not) was randomised.

3.2   Earlier Interpretations of Punishment Results

To understand the full literature and history of work in costly punishment, it must first be understood that one of the 'holy grails' of anthropology is explaining human uniqueness. Why are humans the only species with advanced technology? Why are we dominating the biomass of the planet with our ever-expanding population? The explanation cannot be entirely biological, because the vast majority of technological complexity is of very recent origin, given that very human-like species existed and used primitive tools for millions of years. Urbanisation, agriculture, writing and doctrinal religion (religions shared outside close-knit tribal structures) all seem to date to no more than 8,000-12,000 years ago, well after the appearance of *Homo sapiens*.

Numerous empirical and theoretical studies have emphasised the human propensity to cooperate as a possible explanation for our the presence of culture (e.g. Gintis et al., 2003; Henrich et al., 2001). However, this only presents a chicken-and-eggs problem — is this an explanation or a rediscription? What accounts for this level of cooperation? After the early PGG results (e.g. Fehr and Gächter, 2000; Fehr and Gächter, 2002), altruistic punishment was originally regarded as a possible explanation for cooperation. Here too though the reasoning seemed cyclic, as punishment was a form of altruism itself. Due to previous research on human cultural evolution I had become part of a cohort of researchers that did not find it likely that altruistic behaviour either difficult to evolve nor uniquely human (Čače and Bryson, 2007; West et al., 2007; Bryson, 2009). This was why the Herrmann et al. (2008a) results originally appealed to me — they indicated that punishment was part of a much more complex system of social regulation rather than a simple explanation for human culture. More recently, the phenomenon of ASP has lead some scientists to emphasise the 'dark side' of human behaviour, including a tendency for spite and hyper-competitiveness (Abbink and Sadrieh, 2009; Jensen, 2010). Swings of moral assessment and defensiveness need to be guarded against if we are to understand what underlies these phenomenon. The research presented here attempts to take an objective, biological perspective: existing behaviour needs to be described and explained — where did it come from, and why has it been sufficiently successful to persist?

3.3 Proximate and Ultimate Explanations

Finally, a brief introduction to some basic biological nomenclature. Although like much in evolutionary biology the exact concepts and terms presented here are still subject to debate and refinement, but there is broad agreement on their general meaning. Any behaviour we see in nature is generally expected to have a number of causes. *Ultimate causes* concern why the behaviour is present in a population as whole — what role does it serve in the evolutionary struggle? Contemporary biology does not expect all traits to be adaptations — some are incidental side-effects of historical associations, since the selection process takes time and can only operate on the material at hand. A *Proximate cause* is essentially a mechanism, what makes the currently-observed animal do what it is doing. So for example running may be ultimately a good way to escape, and proximately caused by a loud noise. Note that for some species, flying or swimming is a better mode of escape than running — this has to do with their evolutionary history. Just because something is an ultimate cause of a behaviour does not mean that behaviour is necessarily the optimal mechanism for meeting that need.

Note also that a useful proximate mechanism may itself become an ultimate explanation for some other trait. For example, the utility of running from some noises but not others may result in selective pressure for being able to discriminate those noises.

4 Scientific Outcomes Concerning Antisocial Punishment

4.1 A Negative Outcome: Automatic Clustering

One of several theories entertained in our proposal was that different societies were composed of different types of individuals, or perhaps different proportions of different types. By 'type' here I mean primarily 'plays a recognisable strategy across the set of PGGs', that is that we were looking for categories of long-term strategies over the 20 games. With HSP, we had hoped to see if these strategies correlated with personality traits. One could conceivably check for other genetic or cultural/environmental correlates (e.g. birth order, affluence), but those were never in the remit of this study.

However, the postdoctoral assistant involved (Pablo Lucas) was never able to find a reliable automatic clustering system that worked, despite attempting a wide range. Of course a negative outcome for a search is not in itself a proof that no possible solution exists. We ultimately developed some very simple heuristics for clustering strategies by hand, which are discussed below with their related outcomes.

4.2 The Terminology Behind ASP and AP Is Misleading

The next several results are due primarily to the work of Karolina Sylwester though in collaboration with James Mitchell and myself, and with assistance from Lucas, Lowe and Powers. Publications about these are those submitted with Sylwester as first author.

One thing we did find from examining the data is that the terminology behind ASP and AP are quite misleading. 'Altruistic' punishment is only altruistic ultimately if at all. Proximately all costly punishment seems frequently motivated by aggressive tendencies. Indeed, as we will discuss below, punishment may also have consequences in establishing social dominance, whether or not that is its intended purpose. Secondly, ASP is not always aimed at top contributors or in response to revenge. Sometimes it is aimed from the lowest contributor to the second-lowest contributor, in an apparent effort to make them produce more public goods while allowing the punisher to continue to free ride. Speaking strictly biologically, this can actually be seen as an altruistic act, because the punisher had to pay a penalty, and the other members of the group who are not the punishee benefit just as much as the punisher if the punishee increases their contribution.

The fact the terminology is misleading does not mean it should be abandoned. ASP and AP both has clear definitions and clear correlates with important measures of economic well being. We will show more that we discovered below. But this is a reminder that the terms cannot be used for obvious moralistic assessment and that socio-economic behaviour and dependencies are highly complex, particularly at the population level.

4.3 Difference in Propensity for ASP Seems Correlated with Out-Group Assessment

One of the promised contributions promised in the first year of our grant was a review of all literature relevant to ASP. We expanded this to look at costly punishment in general, because as just implied, our preliminary research indicated that sometimes the labelling between AP and ASP are deceptive, and it is not clear that ASP can be understood independently of systems containing AP as well.

Our basic finding was that because experimental subjects used to investigate costly punishment originally had been heavily biased towards participants from democratic and relatively wealthy places (Henrich et al., 2010), antisocial punishment has, historically, been regarded as the "ugly step-sister" to altruistic punishment and treated as a rare phenomenon, not deserving of scientific attention. Our review gave evidence that the contexts where antisocial punishment is pervasive seem to be the ones in which being *locally* competitive (competing for dominance with nearby neighbours) provides a considerable improvement in the socio-economic condition of the individual. In these contexts, cooperation actually remains more stable even in the non-punishment condition (see Figure 2) but at a lower relative level than more globally-competitive regions are able to achieve when punishment is produced. This appear to be because a small fraction of individuals exhibit a preference for competition, though see section 4.4 below.

This implies that antisocial punishment does not have to be viewed as an exceptionally complex or perplexing behaviour. Rather, it can easily be perceived as simple aggression driven by competition.

One way to think about local vs. global (in the biological sense) competition is that local competition occurs *within* groups — e.g. who in my family gets the biggest piece of pie? While global competition occurs *between* groups — e.g. which family gets the most pies? This is related to the zero-sum assumption — if there is only one pie of a fixed size, then the size of my piece matters a lot (zero-sum). If there is a way to create more and bigger pies, then it could very well be worth collaborating to do so.

Our review article describes theoretical reasons to expect that AP would be less useful in between-group conflicts, since it might prompt other groups to behave more competitively. We also review evidence — scant but promising at the time — that ASP is practiced less in contexts where the other group members are assessed as "in group". This evidence has been supplemented very recently by Lamba and Mace (2011), who show that in extremely similar but discrete populations of a very small-scale minority culture in India (the Pahari Korwa), there was less ASP in villages that contained more of other cultures compared to pure Pahari Korwa villages. This indicates that the presence of a potential out group made a game played between members of a single culture seem like an in-group games, but where obvious out-groups were non existant in the broader population, subjects viewed their co-culture-members as potential competitors rather than collaborators.

4.4 Variation in ASP Correlates Best with Difference in Proportions of Highly Cooperative Actors

I now return to the question of types of individuals or strategies extant in populations as an at least proximate explanation of ASP. One might expect that ASP would lead directly to reduced contributions just as AP leads to increased, but in fact victims response to ASP was much less directed
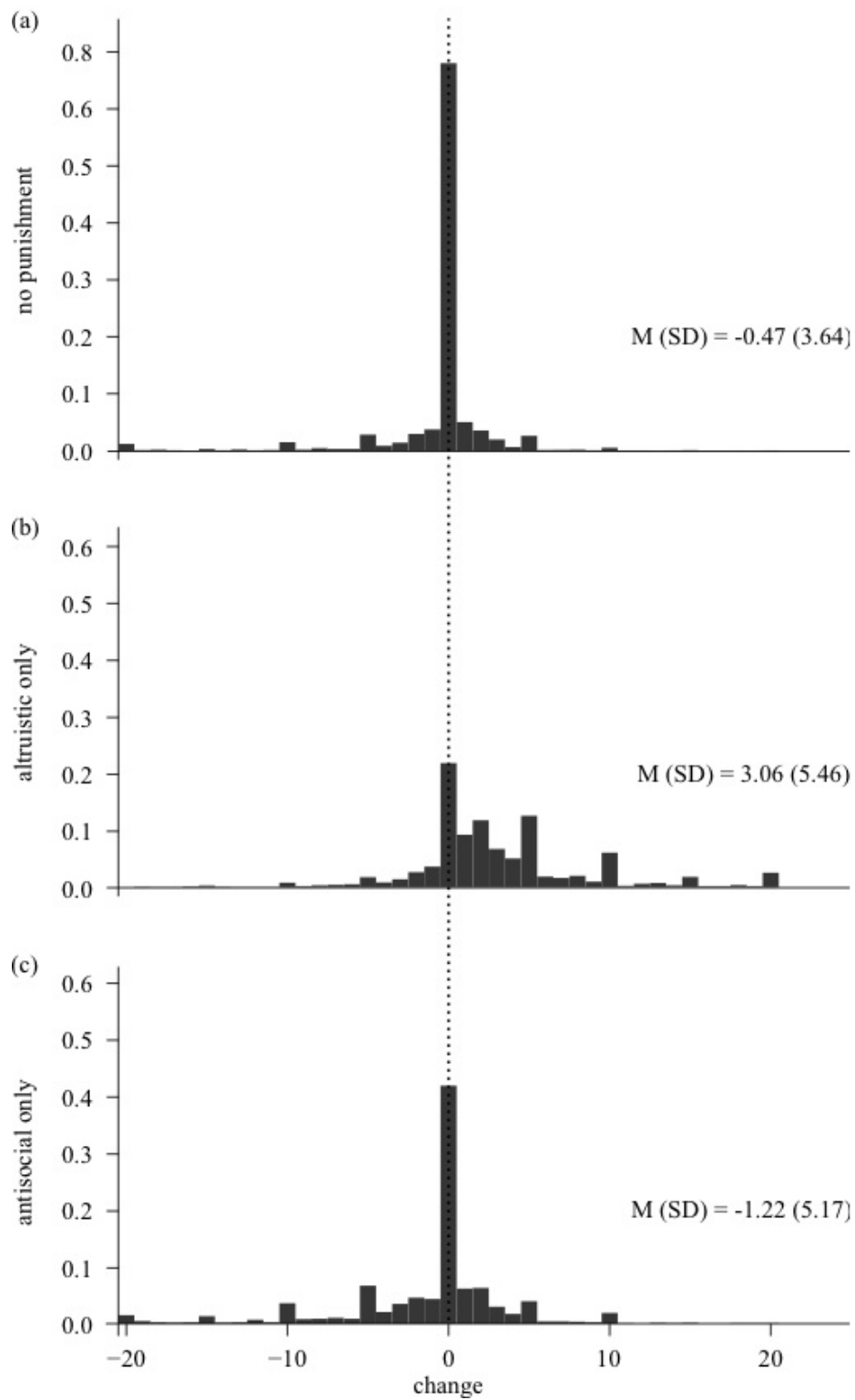
Figure 3: Histograms showing the change in contributions between round n and n+1 after experiencing different types of punishment.
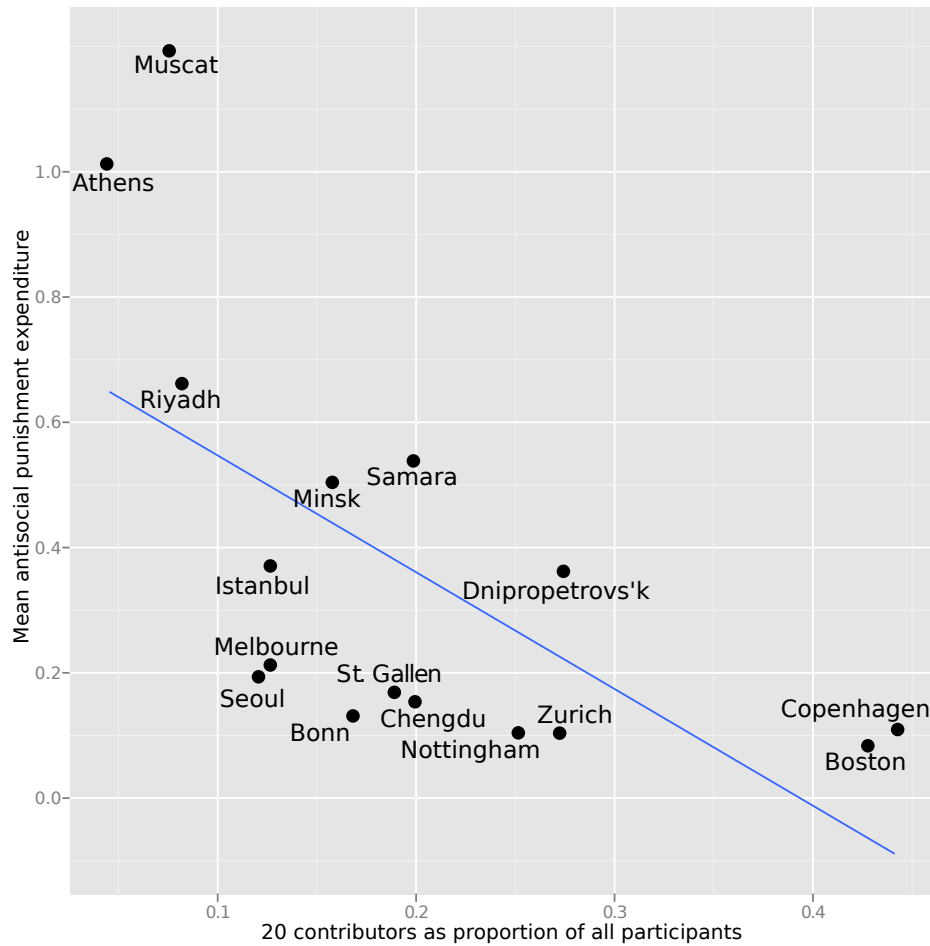
Figure 4: Subject pools plotted by mean amount of ASP (y axis) and the proportion of subjects who contributed all of their available resources (20 tokens) in the first round.

than victim response to AP (see Fig 3) . To investigate what did actually correlate with decreased contributions we explored, the hypothesis that subject pools might differ in the composition of cooperative types i.e. people predisposed to cooperate or to exploit others. For clarity (and after some experimentation), we focussed on distinguishing just two classes of extreme behavioural types. Our classification was based on participants behaviour in the very first round of the first public goods game they played, in cases where no punishment was allowed. Recall that subjects will have read the instructions and demonstrated their comprehension in a test — the first move therefore signals their their interpretation of likely events. After the first round, we tend to see a good deal of conformity bias, that extreme contributors tend to move more towards the group average, but still maintain a bias towards their initial action.

We classified those who invested their entire initial allocation to the group account as 'Cooperators', while those with who did not make any group investment at all, as exploitative 'Free-riders'. The rest (the vast majority of participants) we did not classify. We reasoned that if a person devotes their whole allocation to the group welfare, full cooperation is likely their default behaviour when interacting with strangers. Analogously, we assumed that people who do not make any effort to support their new group have a tendency to behave in an exploitative fashion, or at least not to trust others to cooperate.
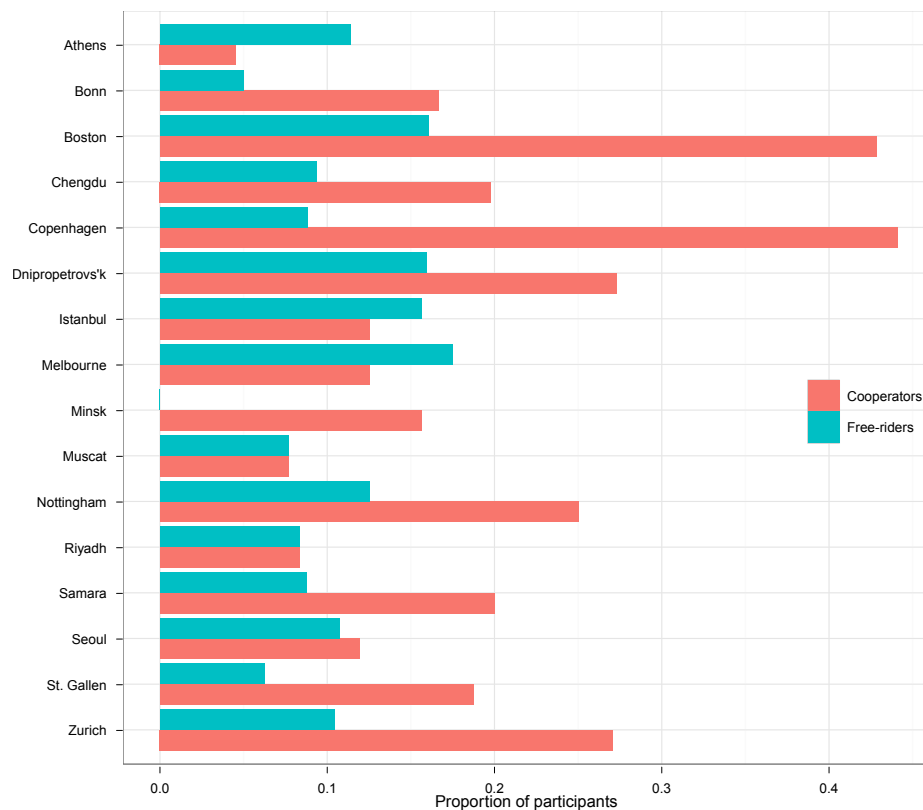
Figure 5: Proportion of participants that contributed all (Cooperators) or nothing (Free riders) in the population. Sorry, this graph has subject pools in alphabetical order, not ranked by ASP.

We found that the variation across subject pools in the proportion of Cooperators is much greater than the variation in the proportion of Free-riders (see Figure 4). We then ran correlations to determine whether there is a link between the proportion of cooperative types in a subject pool and the mean expenditure on antisocial punishment. There was no correlation between AP and the proportion of Cooperators,between AP and Free-riders nor between the proportion of Free-riders and ASP However, we found a strong anticorrelation between the proportion of Cooperators and ASP ((Figure 5).

This means contrary to expectation that the variation between cultures may be primarily the difference between the probability of individuals playing an optimistic, cooperative strategy. Such behaviour may actually *inhibit* the expression of ASP rather than trigger it, perhaps by signalling in-group affiliation.

4.5 Ultimately, ASP is Not Viable Unless It Correlates with Some Other Benefit

I now turn to ASP scientific results that were found through ABM. Like all ABM, these results are theoretical in nature, meaning that they show what is possible in the world if our assumptions about it are correct. The general utility of such results is discussed further below in Section 5.1. Here I take this as an accepted scientific standard and proceed with the results.

These experiments were run primarily by Simon Powers though in daily collaboration with Daniel Taylor and with assistance and direction by me, and some assistance on understanding the empirical data by Sylwester. Most of these results are available in the paper Powers first authored,
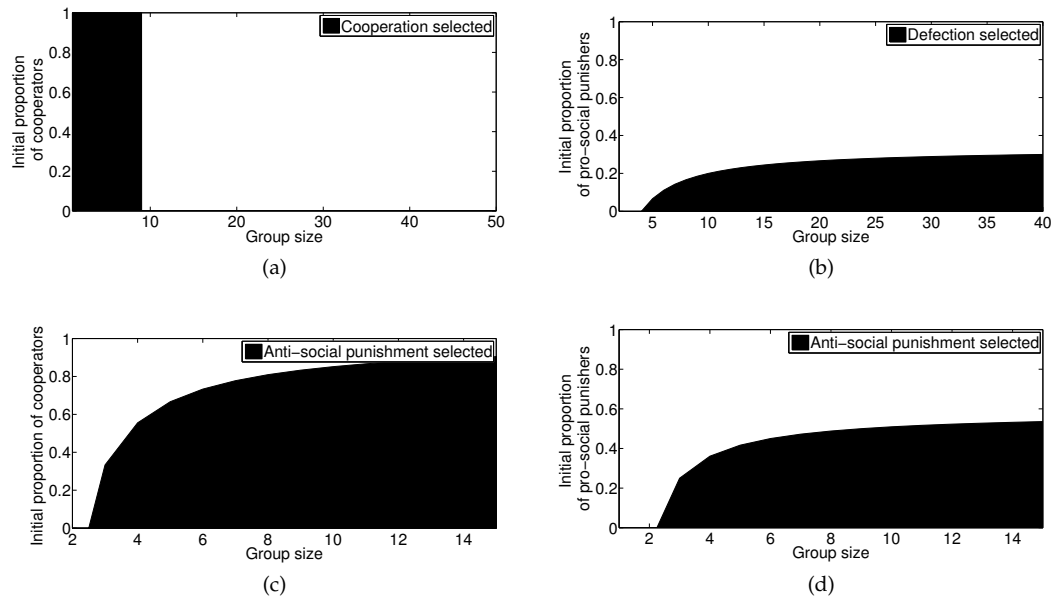
Figure 6: Formal analytical (not simulation) results of pairwise competition between strategies. Groups are formed by binomial sampling from the global population and disperse every generation (trait-group model). Parameters are as used in the simulations (B=0.9, C=0.1, K=0.1, P=0.5). A) Cooperators vs. defectors. B) Pro-social punishers vs. defectors. C) Anti-social punishers vs. cooperators. D) Anti-social vs. pro-social punishers.

but some are not yet written into a paper and have only been presented as talks.

Based on the review of literature (which was advanced when Powers joined the project) the ABM we ran were not spatial as originally expected, but rather focussed on global vs. local competition. This is to some extent a standard framework for studying cooperation. However, Powers 2010 PhD contribution (recently reviewed in *Science* by Szathmáry, 2011) was to extend this framework to consider the *size* of the group as well as the extent of within and between group competition. In a context where there *is* some limit on the ultimate advantage of cooperation, small groups may benefit more since the advantage of the cooperation (for example, being able to keep prey by driving off scavengers) can be shared between a smaller number of individuals, giving each a greater reward.

This situation is similar to the one found by (MacLean et al., 2010) in the production of public goods by bacteria. Where that public good is the production of enzymes to digest the food the bacteria are situated in (just one of many public goods bacteria produce), it only makes sense for the bacteria to pay the cost to produce that enzyme up to the level at which all the food is digested. MacLean et al. show that rather than evolving exactly the correct level of production (which could be difficult particularly if that level changes quickly), a species bifurcates into altruists and free-riders at an appropriate mix. Achieving the right mix in bacteria can be achieved with simple mortality, since their lives are relatively cheap even from a species-level perspective. If anything similar can explain the variety of strategies we see in humans, then given the high cost and long lifespan of humans it is far better if switches in strategy are achieved by psychological factors, such as the in-group/out-group assessment suggested in Section 4.3.

MacLean et al. show that such mixed strategies are only possible to evolve if the altruists get at least a slight advantage. In the case of the bacteria, enzyme producers get slightly more access to food. Our own simulations showed something similar for ASP. As explained in Section 4.2, ASP is
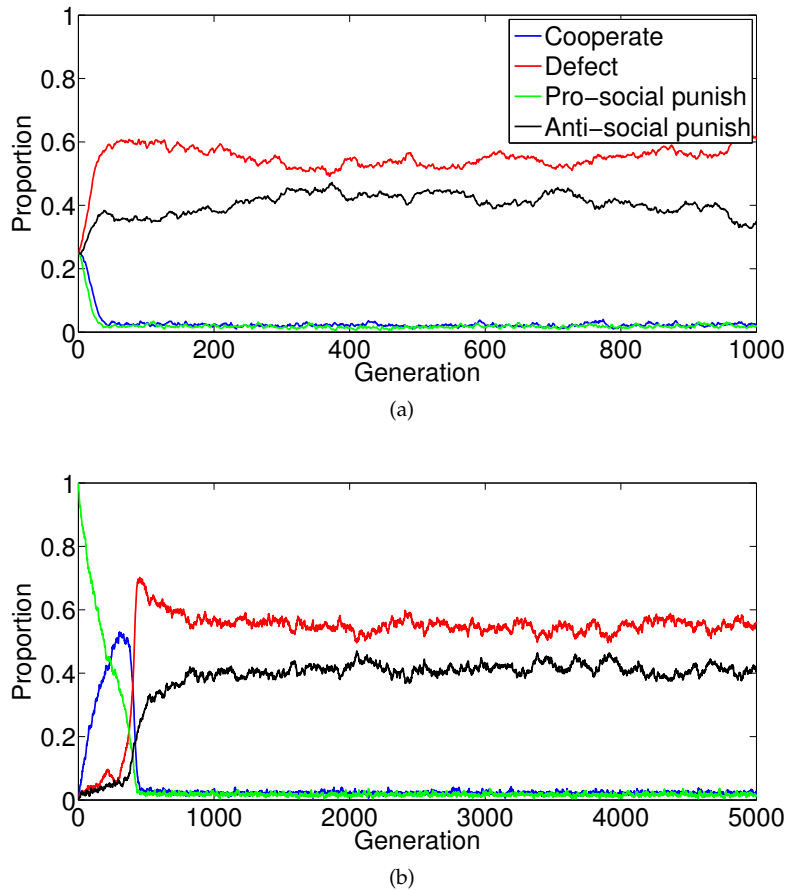
Figure 7: Simulation results of competition between pro-social punishers, anti-social punishers, non-punishing cooperators, and non-punishing defectors in a well-mixed population with mutation. A) All types initially present in equal frequency, showing the selective advantage of defection from this initial frequency. B) Pro-social punishment initially at fixation. The increase in non-punishing cooperator mutants creates a second-order Tragedy of the Commons in which first pro-social punishment, and subsequently cooperation, collapses.

actually costly and, relative to other group members, altruistic, since they too benefit by the loss of relative fitness of the punishee, yet pay no costs themselves. We could find no evolutionary context in which ASP was adaptive against other social strategies unless we gave ASP a small energy advantage. What could this advantage be? Our best guess at this point is that ASP, indeed costly punishment generally, has a side effect of giving those that express it social dominance status. This in turn may result in long-term payoffs in getting better access to resources at lower cost, since in general dominance can be viewed as a form of long-term conflict resolution (Preuschoft and van Schaik, 2000; Bryson et al., 2012).

4.6   ASP Can Promote Defection in Group-Structured Populations

A related but likely to be more disruptive finding is that we have shown that the presence of individuals who play ASP strategies in a population completely undermines that very common assumption outlined in Section 3.2, that punishment explains humanities exceptional levels of co-

operation or even necessarily promotes it. The effect of punishment is complex and depends on context such as the extent of within-group vs. between-group competition as well as group size (see Figures 6 and 7 for some details).

The general scientific result then is that at an ultimate level, punishment can be used *either* to increase or decrease cooperation, and that depending on the socio-economic context *either of these might be a good move*. There can be a context where too much is invested in public goods. It must be remembered though that every population studied has shown positive levels of altruistic cooperation. The only question is how much is expressed, and what indicators control this level in human populations. This will be discussed further in the next section, on policy.

5   Policy Implications of ASP Research

The assumption of the scientific aspect of our project is the standard one made in biology: that the seemingly bizarre behaviour of antisocial punishment is a side-effect of strategies that are generally advantageous — or at least not disadvantageous — to people living in some cultural contexts, presumably the ones in which it is found. Our goal for the three-year project was to identify the conditions in which the expression of antisocial punishment might be encouraged in a culture, and the conditions in which it tends to be extinguished. In this first year of study, we were able to show that:

1. ASP is a disruptive more than a reliably "down-regulating" influence on cooperation. It does not reduce cooperation as reliably as AP increases it.

2. Down regulating cooperation might make sense for an individual if that individual's well-being is determined more by local competition (e.g. who is most dominant in a household, village or business) than by global competition (e.g. which household, village or business does best.)

3. ASP seems to be more likely to be expressed in contexts where group members do not expect by default that other group members are "in group". With respect to the previous point, this implies that there is always *some* cohort of trusted individuals, the question is how large it is by default. In Northern Europe (and Boston, the only US city surveyed here), it seems by default to encompass group sizes at least as large as a single university, while in Greece, Turkey, the Middle East and the former Soviet Union it does not.

4. Whatever the *default'* level of in-group assessment is, some manipulations can alter this. The only ones we could explore without doing human subject experiments[2] were the natural experiments of seeing how subjects respond to having someone in their group who contributes *all* of their resources to the public good, and of having someone in the group who contributes *none*. Interestingly, we have learned here that having super-defectors in the group *has no effect*, but having super-cooperators in the group *reduces ASP*. This implies that people inclined to ASP are impressed by such a clear expression of in-group assessment, and have some tendency to believe it and adopt it.

The final point may sound promising from a policy perspective. Note though that in the experimental context, commitment of resources is completely transparent. All subjects know they have equal access to information and equal power under the authority of the experimenter. In a more realistic context, showing total economic commitment or some other signal of in group affiliation may be difficult to make convincing.

---

[2]because we never achieved HSP from the AFOSR SGEC, despite achieving it several times from our local IRB, see Section 2.
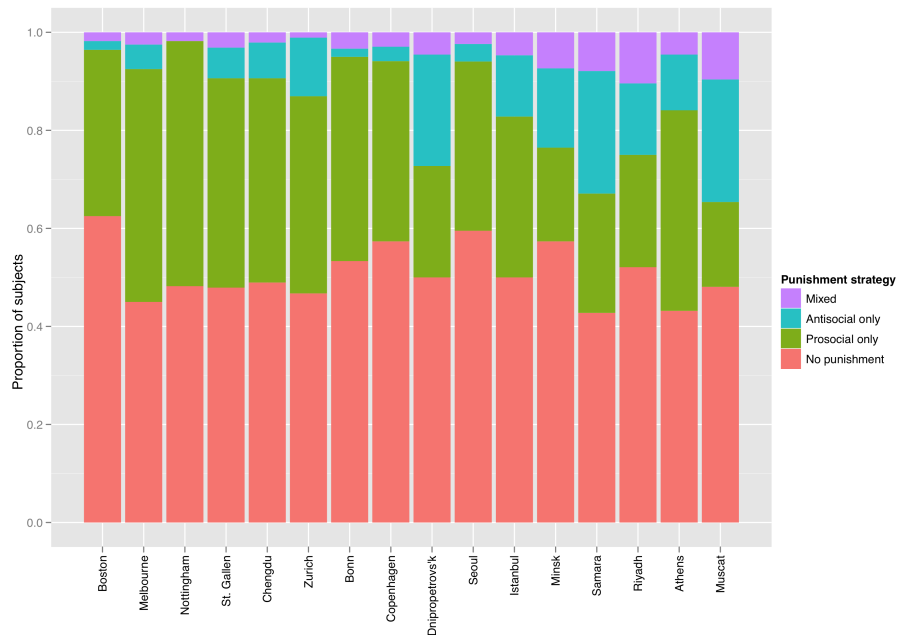
Figure 8: Proportion of participants in subject pools that used various punishment strategies at all in the first round of play (before being punished themselves.) 'Mixed' means the punishment of more than one other player, at least one of whom contributed more and at least one of whom contributed less. ASP and AP mean *only* that strategy was used. Note these figure do not show the *extent* of the investment in investment so does not allign perfectly with overall ASP investment

Certainly as Americans we can identify with the "in group" assessment that comes from knowing someone else has chosen the same college or university as we have, particularly in the same or similar year. College is a significant investment — even where tuition is free, college requires 3–5 years of a person's life. For us, making similar investments at this scale is enough to incline us towards in-group trust, but then we live in societies with a high Rule of Law (c.f. Section 3). Understanding the social experience of those who cannot make this assumption about their college colleagues is work for us. Most of us will have had *some* experience of being in some situation where we were not sure everyone in the room was interested in collaborating for our mutual common good — where we have felt in danger of exploitation. The point is that in some cultures that feeling appears to extend even to the prestigious university campuses that Herrmann et al. (2008a) chose to study[3]. This might indicate that it could be difficult to achieve trust and therefore high levels of economic cooperation in other professional contexts as well as university.

We must remember that in every society studied, ASP practiced by some participants, but similarly in no society was it practiced by all. Figure 8 shows per city (in the order of ASP pervasiveness) the proportion of subjects that use ASP either together with AP (top, purple colour) or exclusively (second from top, blue colour) *in the first round.* (Note that this also proves that ASP is not just about revenge for AP, another early theory — no punishment has been experienced yet.) Figure 9 shows rounds 2–9[4]. One of the original goals of our study which could not be realised without HSP was to find out if we could identify in advance personality indicators for predisposition to ASP. These may

---

[3]Because the initial studies were conducted at ETH, it was considered essential that representatives from other cultures were also drawn from top universities to increase comparability.

[4]Round 10 is excluded because it is also exceptional due to having no following round. Levels of ASP are similar there as for round 1; levels of AP drop but do not disappear.
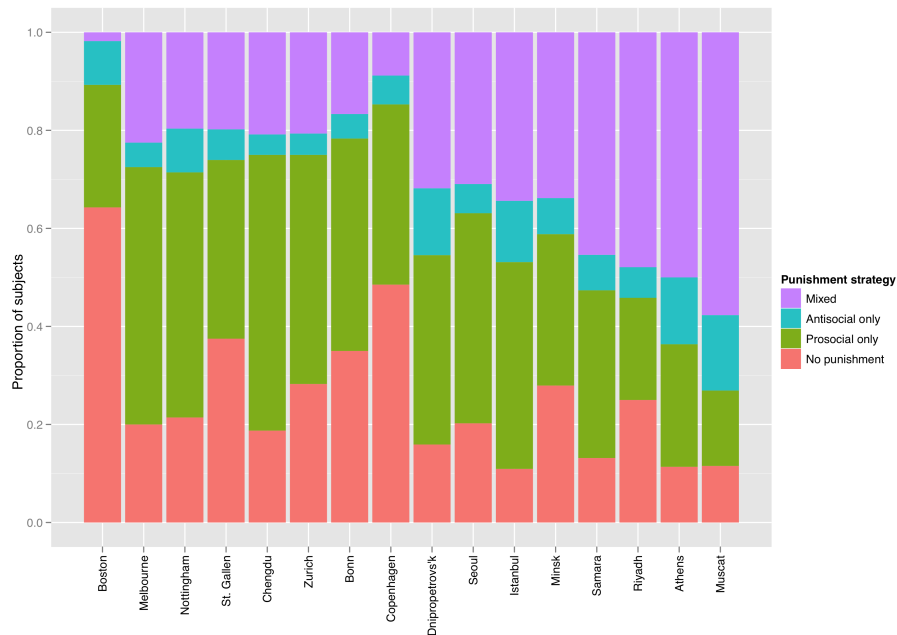
Figure 9: Proportion of participants that *in any round* used various punishment strategies in the middle rounds of play (2–9)

not exist. ASP may respond primarily to a combination of present and cultural context, combined with an element of stocasticity. However, even if we could determine who practices ASP, we have no idea of what the broader impact for a society would be if these individuals were excluded from positions of power or negotiation. As we mentioned, in some circumstances reducing group size or down-regulating public investment may make economic sense.

We also do not know for sure that decreasing ASP and/or increasing cooperation would increase GDP. It could well work in reverse — where individuals are affluent they can take more risks about in-group inclusiveness. It seems likely though to be a situation of mutual feedback, and that if honest, transparent signals of mutuality of interest can be established, higher levels of both cooperation and economic performance could be established.

5.1   AFOSR Policy Towards ABM

I would like to speak briefly to the original goal of Terry Lyons that motivated this entire project: the ability to increase the standard of scientific practice in ABM. First, it is unambiguous that some people already practice ABM in the highest-standard of science. This is unambiguous because scientific standard is established by peer review and citation, and there have been ABM papers in the highest-impact scientific journals for almost a decade now (e.g. Azevedo et al., 2006; Powell et al., 2009; Bowles, 2009; Powers et al., 2011) and this trend is increasing.

At the same time, it cannot be denied that the majority of ABM papers are not in the pages of *Science* or *Nature*, but then neither is the majority of behavioural economics. The primary problem for the AFOSR and all funding agencies is discriminating promising projects from projects with less promise. Excluding all ABM work is not likely the best solution to this, but being more selective may be. In my opinion, the original goal for this project — of doing methodological work on helping the average modeller seeking to improve their standard of science — is still a valid and

promising one.

The nature of science is that much unexpected is discovered in actually performing the experiments — experiments seldom either validate or invalidate hypotheses, but more often lead to unexpected outcomes. ABM and simulation more broadly contribute to the scientific process both through an initial process of specifying a theory in code, which can lead to immediate discovery of incoherence or incompleteness where present. If successfully completed, a model provides valuable evidence of theoretical coherence. It can sometimes also lead to discovery of unexpected side effects of a theory which are therefore predictions to be tested against data. This will be discussed further below.

6 Scientific and Technical Outcomes Concerning ABM

As just mentioned, this project also had a strong methodological component. Our proposal was originally solicited with the aim of improving the current state of the art for incorporating simulation into experimental social science, particularly in anthropological research. In our proposal, we promised:

> To this end we will extend — and, where necessary, develop — tools, analytics and protocols for increasing the interplay between experimental economics and social simulation. Our goal is to accelerate scientific discovery and deepen understanding. We will achieve this initially by facilitating the communication of between the modeler and the experimentalist, and ultimately by making modelling a standard mechanism for communicating scientific theories.

In this section we review our theoretical contributions to this problem, our formal research outcomes, and finally provide an informal description of our experience.

6.1 The Role of ABM in Social Sciences

In response to my initial conversations with Terrence Lyons, and in the context of the Oxford *Explaining Religion* project, I wrote the following (taken from the introduction of the Whitehouse paper listed in publications, below):

> Since its introduction to the social sciences by Axelrod and Hamilton (1981), social simulation has been widely viewed as a powerful tool with obvious applicability into these disciplines. Nevertheless its history has been spotty. There have been outstanding recent examples of social science contributions informed by simulations such as the Powell et al. (2009) anthropological account of the Palaeolithic transition, or the work done by Laver (2005) to test the fit between theories of political party policy making and real party behaviour. Nevertheless the vast majority of social simulation papers languish in specialist conferences and journals. While such venues may help advance the technological state of the art for simulation, the substance-area contributions of these works are overlooked and indeed possibly untested by the review processes outside of the mainstream for their application areas. This process may account for the suspicion with which some academics view work in social simulation.
>
> Of course all new research methodologies, whether quantitative or qualitative, meet deserved resistance before they have proved their utility. One aim of the present article is to argue that social simulation has by now proved its utility, at least in some contexts, and to help make those contexts more apparent — not only to those who do not yet include social simulation as part of their research tool set, but also to those modellers who

have had difficulty publishing in or otherwise engaging with the mainstream literature of their chosen field.

Although there has been some debate and discussion about the exact role of simulation in not only social science but science more generally, there is increasing consensus that a simulation simply *is* a theory (Bryson et al., 2007; Kokko, 2007). Like any other theory, a simulation must be tested against data derived by observations of the world — of the target system that the theory seeks to explain. What makes simulation special as a methodology is that in order to be simulated, a theory must be specified so completely that its consequences can be demonstrated by executing it on a computer.

The output of simulations once run should not be mistaken for data about the world or about the target system. Rather, the output of a simulation is data about the *theory* the simulation embodies. As Kokko (2007) puts it, we need models because "our brains aren't big enough" to understand all the consequences of our theories. The output of a simulation can be tested against data derived from observation using standard techniques of hypothesis testing. Theories that account for data better than other theories become dominant, where "better than" can be either with a closer fit or with fewer assumptions or a more general application.

Importantly, theories can be modified. They can be extended, generalised, or specified to create more or better explanations about particular systems they might be applied to. Science itself can be viewed as an evolutionary process, where peer review and other forms of expert opinion constantly narrows the field of available theories by selecting between them, but then subsequently broadens the field by applying a variety of enhancements to the theories in order to increase their validity. Simulation can perform several key roles in this process. It allows theories to be communicated with unprecedented fidelity as the actual code of the simulation can be transmitted from researcher to researcher over the Internet. It allows the theory to be modified through modifying the code, and for the consequences to be checked and validated against data by executing the code.

The simulation process therefore promises to accelerate the scientific process, though it also carries with it a risk. It is possible that theories represented as computer simulations could become so complex that they could not be represented or truly understood by any one mind. This could leave us with a situation where our ability to *predict* behaviour could continue to improve, but without carrying with it increased human *understanding*. Besides the fact that this may decrease the pleasure or accessibility of scientific reasoning, it might also result in a reduction of scientific innovation where that innovation is based on human insight.

However, the problem of theoretical obscurity is not unique to computer modelling. Many scientists and philosophers with no connection to computer modelling have proposed complex theories, or even simple theories about such complex entities that the consequences of those theories are not clear. This brings us to another domain of utility for social simulation. It is frequently reported anecdotally that one of the advantages of adding a modeller to an existing research team is that modellers tend to ask a different set of questions. This might be seen as a normal and expected consequence of any interdisciplinary research, and that may indeed be part of the explanation. Where a modeller has been trained for example in the computer sciences they will bring a different educational background and perspective to a social science problem. But the contribution of simulations goes beyond that. Quite simply, expressing a theory in such a way that it can be run on a computer requires a completeness and specificity that is often beyond the ordinary requirements of academic publishing. Thus building a working social simulation may actually require a great deal of theory construction.

University of Bath, Department of Computer Science   *J J Bryson, PI*

In this article we seek to demonstrate that this process of theory construction and communication is in itself an academic contribution. The best-established benefit of the simulation process is that it allows the consequences of the theory to be better understood and therefore opens that theory up to refutation, validation or improvement after comparison to observed data. But the importance to a science of simply forcing the theoretician to more complete in describing their their theory should not be underestimated.

The article goes on to show how Prof. Harvey Whitehouse came to better understand and further develop his well-known and well-established modes of religiosity theory due to the experience of building two different ABM, one that illustrated the formation of a schism, and the other that explores the cycling between fanaticism and stability that his theory predicts.

6.2 Additional Scientific Outcomes Beyond ASP Results

Many of the scientific outcomes previously reported under ASP (Section 4) were derived from simulations and formal theoretical analysis, starting from Section 4.5 (p. 14) above. Here we briefly review other published or submitted outcomes relating to more basic questions of cultural evolution.

6.2.1 The Size of Culture

Čače and Bryson (2007) demonstrated that there is selective pressure for sharing information as a public good. As explained in Section 3.2, this sort of result begs the question of why humans have uniquely large simulations. By extending both the simulation and the analysis, we have come to realise that what limits the size of culture is the amount of information that can be reliably transmitted from one generation to the next. The size of the culture is therefore affected by things like population density (how many people can you learn from at any one time?), lifespan (how long can you learn?), rate of transmission of information, reliability of transmission, and memory capacity. A paper on this result is under revision by Bryson, Lowe, Bilovich and Čače.

Two important consequences of this result are: given that the amount of information is determined by ecological / biological constraints, the *quality* of the information is at least as important. This is congruent with the fact that IQs continue to increase far faster than biological evolution could explain (Flynn, 1987). Much of our intelligence is a result of the quality of the information we choose to transmit, and that too is constantly improving.

The second consequence is more surprising — it is possible that, just as with levels of cooperation there are limits to the ideal size of culture. It is possible that limits to quality of memory have the effect of limiting the size of culture, but also its persistence, which may allow for greater innovation or the ability to track economic or other environmental change more easily.

6.2.2 The Persistence of Culture

This latter consequence was derived not only from the previous study, but also from another one originally designed to examine the impact of modularity of mind on cultural stability as proposed by Sperber and Hirschfeld (2004). This second simulation was originally presented at a workshop by Bryson (2008) on cultural variation, attended following the submission of the grant proposal to the AFOSR and due to conversations with Lyons. The aspect of the simulation relevant to the present context is that it showed how remarkably persistent culture is in the face of even very small amounts of bias toward conformity with ones neighbours. It also showed how the disruption caused to a culture because one aspect of that culture's beliefs are replaces with a truly significant improvement can lead to a general diversification of subcultures that become relatively persistent

again once the selective event stabilises. This work is now in press as a book chapter, but bears further analysis and relation to anthropological data.

6.2.3   Culture Can Benefit Any Species

One theoretical impediment to our research assumption set out at the beginning of Section 5 — that culture can be studied in the context of assumptions rooted in biology — is something called *The Rogers Paradox* (Rogers, 1988). Rogers thought he had shown that culture is always a neutral adaptation, because it will only persist in a population if those that use it as an information source get no advantage over those necessary individuals that first derive the information. This paradox is much discussed for obvious reasons, and again makes human culture sound weird and inexplicable (Boyd et al., 2011; Lehmann et al., 2010; Boyd and Richerson, 1995, 1996; Rendell et al., 2009; Richerson and Boyd, 2005; Laland, 2004; Enquist et al., 2007; Henrich and McElreath, 2003). Taylor with Bryson showed that the apparent paradox is a consequence of confounding two different definitions of culture. This work is submitted for publication.

6.2.4   Primate Societies and Genetic Evolution

As discussed just earlier, one consequence of being in a species that depends heavily on culture as a source of behaviour is that we need to have long lifespans. This seems to be true of cognitive strategies in general, and we share our long lifespans with other species such as many of the non-human primates, particularly monkeys and apes. The problem with longer lifespans (particularly including longer periods of childhood) is that they reduce the rate at which biological evolution can be achieved, since childbirths are fewer and further between. This can have negative consequences for things for which biological adaptation might be useful, such as immunity or climate change. Yet the rate of speciation (a good indicator for evolution) in most mammals is an order of magnitude faster than in most other vertebrates, and faster still in social mammals (Di Fiore and Rendall, 1994; Baker and Marler, 1980). The Primate order is amongst the fastest-speciating orders of mammals. The most speciose family of Primates, and also the youngest, is the Cercopithecidae (Purvis, 1995; Perelman et al., 2011).

One of Bryson's persistent interests has been in the extent to which cognition or cognitive strategies might have knock-on effects for accelerating biological evolution, as first hypothesised in the nineteenth century e.g. by Baldwin (1896); Hinton and Nowlan (1987). Gluckman (a self-funding PhD student) looked at a particular aspect of primate social structure, matrilineal organisation, to see how it affected the rate of evolution. He discovered a clear pattern of acceleration of selection directly proportional to the strictness of the dominance hierarchy in a species. This makes sense since dominance ordinarily is considered to covary with fitness both in innate biological predisposition and in expressed fecundity due to extra resources being available to mothers. Nevertheless, we know that the strictness of dominance rank is dependent also on other biological and ecological factors. Again, this result is part of an ongoing enquiry, but is another indication that culture is not necessarily a handicap.

6.3   Related Outcomes

Research in this section was related to this project, but could not be costed to this project nor acknowledge the AFOSR because the experiments required human subjects. Nevertheless, one set of outcomes in particular bares on future funding. Michael Tremante in his undergraduate final-year dissertation explored the use of Second Life, and on-line virtual reality "world" used globally, for PGG experiments. Some of his results were promising, but of two low a number of subjects to be considered scientific outcomes. However, viewed as a pilot study, his research did show that there were substantial issues with the approach. Subjects found in Second Life were much less likely to be

able to pass the training and exams, and also would often interrupt the experiment due to personal events (dinner, a phone call). It was also difficult to ensure anonymity in a medium where most avatars are designed to have life histories, friendships and communication. Thus while virtual reality may be promising venue for manipulating in-group/out-group indicators or maintaining consistency across regions, it is not viable in itself as an experimental platform. Rather, subjects should be brought into the controlled setting of a laboratory and given a short-lived avatar. Unfortunately, this produces another set of hurdles as subjects must learn not only the PGG but also at least the basics of Second Life.

The other paper listed in the deliverables concerned programability of human-like agents. The results were inconclusive, but a contribution can be seen in attempting to find a metric for this task. The mechanism used was running experiments on programming students, splitting a course between conditions. The outcome was made a part of their mark / grade for the course, but conditions were balanced such that no student was disadvantaged if one approach proved more difficult to use. This research is continuing.

6.4 Software Outcomes

The full three-year project was originally designed to include a substantial software element intended to make it easier to create complex, human-like intelligence for agents in social simulation. The original grant included 12 months of programming time (9 months in the first year, 3 months in the second) for initial development, then additional funds for the remainder of the grant for part-time consulting. This is because frankly software never works as expected when it is first delivered, nor do users ever use it in the expected way, so we had expected the code to be refined and released over the full three-year period of the grant.

Because we were fortunate to hire an exceptional programmer willing to work at academic salary (James Mitchell), and also because of the breaking Hauser scientific crisis, we decided at the beginning of the research phase of the grant that Lucas and Mitchell would also work on producing a replicable research framework specifically designed for supporting the combination of ABM and empirical data such as that derived from anthropologists. This system's working name was the Amoni[5] Reproducible Research System (ARRS), although we anticipate finding a better acronym for it before public release.

The idea of reproducible research is that when one is reading a scientific paper, one should be able to replicate its results exactly given the original data the paper was based on. This has become a standard for some political science journals: all papers have not only their datasets archived in publicly accessible places, but also the R scripts underlying their figures archived as electronic supplements for the article. ARRS is an attempt to do this and go further. ARRS allows entire simulations to be run and statistics to taken and checked in comparison to data sets. ARRS is available for download from that link.

As a side benefit to this work, it also makes it easier for authors to do science as they write science papers. If any model is changed or data is added, the system will automatically update the paper — ideal for revisions whether required by review or by corrections discovered during the writing process.

In practice so far only two papers (Gluckman's and the electronic supplement to Sylwester's submission to PRS-B) have been fully ARRSed. Mitchell's time was curtailed not only because he started a month late and had to leave a month early (with no further funding for maintenance / consulting) but also by Lucas' departure. Mitchell was the only one competent and available to work with Sylwester on the analysis of the empirical data.

The work that was in progress for the ABM editing system (ABODE) was stopped immediately when the second two years of the project were cancelled, though this work is now being progressed

---

5.

by a final-year undergraduate research student. Documentation of the work to date can be found form this ABODE link.

6.5   ABM Practice Outcome

Having modelling and human subject postdocs working together for about nine months provided some direct benefit but still required significant facilitation by the PI, even in the final month and thereafter in writing articles. Although the group as a whole was extremely friendly and socialised together out of hours, both postdocs continued to be surprised by the mutual utility of their research and to require prompting to report and consider results useful to each other. It is possible that this cultural barrier would have been overcome if the project had run to full term and both postdocs had received more exposure to the other's methods and outcomes. Both postdocs were only beginning to return quantified results in the final months we were able to pay them. Most of the articles were submitted while Powers was working for a new employer and Sylwester was unemployed looking for work.

One of the reasons the ABODE project was relegated to lower priority for this project (though it would be very useful for other parts of AmonI research) was because we found in reviewing the literature and working to publish our own papers that making agents complex is seldom good scientific practice. Science is about abstraction — finding relatively simple, important features of the world that can explain a great deal of complexity. Further, science is a cultural product, and as just reviewed in Section 6.2.1 these are limited in the amount of information that can be clearly transmitted. Nevertheless, we hope that with tools such as ARRS and ABODE, and further developments and improvement in the utilisation of ABM, models that account for more complex cognitive phenomena will still be a part of the scientific toolbox of the future.

6.6   Conclusion

> To have received from one, to whom we think ourselves equal, greater benefits than there is hope to requite, disposeth to counterfeit love, but really secret hatred, and puts a man into the estate of a desperate debtor that, in declining the sight of his creditor, tacitly wishes him there where he might never see him more. For benefits oblige; and obligation is thraldom; and unrequitable obligation, perpetual thraldom; which is to one's equal, hateful. But to have received benefits from one whom we acknowledge for superior inclines to love; because the obligation is no new depression: and cheerful acceptation (which men call gratitude) is such an honour done to the obliger as is taken generally for retribution... — Hobbes (1651)

Our work has shown that, as with many things, Hobbes was amazingly prescient concerning the creation of public goods given that he wrote in the seventeenth century, but not entirely right. Our research indicates that anti-social punishment may indeed occur in contexts where other participants are not mutually-acknowledged members of trusted group, yet generosity may in absence of other information be taken as an indication that in fact trust is merited.

We have found that costly punishment is best understood as having impact not only on global economics but also on individual competition, and that the apparently maladaptive behaviour of anti-socially punishing those more generous than ourselves may even in some contexts be a sensible response, when our own well-being is (or at least appears to be) most determined by our relative dominance to our local neighbours, not to how well our neighbourhood performs as a whole. For actors such as the US Air Force that may be more concerned about global than local good, the best course of action is probably promoting the likelihood that the benefits of public goods *are* shared by those who need to cooperate, and ensuring transparency so all parties can be assured this is the case.

University of Bath, Department of Computer Science   *J J Bryson, PI*

6.7   Future Work

We have achieved the primary goal of the project — we have built a better understanding of cultural differences in the expression of antisocial punishment, and more generally in the capacity of societies to exploit economic opportunities. Instead of a theory vacuum, there are now plausible explanations supported by simulation and existing data. Nevertheless, our primary hope in proposing the original three-year project was to discover whether we can manipulate the tendency to express antisocial punishment. This would have obvious utility for all manner of groups including nations concerned with their own economic well being or that of other's. Learning to predict the consequences of manipulations and interventions provides both evidence of true understanding of a phenomena and the potential to help populations address the consequent economic inefficiencies.

Throughout this document we have taken the perspective that the failure to find communal economic optima is fundamentally negative, since it means resources are wasted in conflict and all parties have less access to wealth and its associated well being. Thus the most useful outcome of future research would be to discover social characteristics leading to this failure were relatively easily alterable. Measures available to be taken could have either cognitive (e.g. increased transparency in distribution of economic resources) or emotional (e.g. team building or other stage setting for triggering a state of emotional inclusiveness). If such measures work a societies' citizens or leaders could be trained to recognise and exploit contexts where mutually advantageous outcomes were possible. However, it may be that for some societies such interventions would be impossible, impractical or unethical. Even in such cases, we could at least hope that the outcome of research interventions would still be beneficial. It would help us to at least identify, characterise and possibly come to understand cultures with such differences. This might be useful for selecting strategies in cross-party negotiations, or in choosing between economic policy options or development approaches.

Finally, we have made progress through both tool-building and empirical exploration of a joint project as to how to improve the probability that work done with ABM is good science, not bad. The shift in AFOSR policy with respect to ABM and the loss of over two thirds of the research time for this project meant that this goal too had to be deprecated and curtailed. Ultimately we believe we were pursuing the correct course on this, and we hope we will someday have opportunity to continue.

7 Dissemination

7.1 Talks and Conferences

Talks listed below all mentioned the project and acknowledged AFOSR funding, but only those from the Summer of 2011 could we give talks containing substantial new results of our own. Talks presenting results are marked with **bold-faced** titles.

7.1.1 Reviewed

Note: conference presentations are not listed here if the conference proceedings were archival publications; these are listed below under *Publications*.

1. Simon T. Powers, Christopher Heys and Richard A. Watson "How to Measure Group Selection in Real-World Populations", reviewed talk and paper (but paper did not mention AFOSR) in *ECAL 11 (The Eleventh European Conference on Artificial Life)*, Paris, 8–12 August 2011.

2. Karolina Sylwester, James Mitchell and Joanna J. Bryson, **"When and why do people 'punish' cooperators: Individual strategies and regional variation in anti-social punishment"** *The Twenty Third Annual Human Behavior Evolution Society Conference (HBES)*, 29 June–3 July 2011.

3. Simon T. Powers, Jason Noble, Jordi Arranz and Manuel de Pinedo Garcia "Explaining cooperative groups via social niche construction", reviewed talk and abstract in *Evolution, Cooperation & Rationality 2011*, 27-29 June, 2011, University of Bristol.

4. Joanna J. Bryson, "The interplay between cognition and secondary replicators: Could memetics explain human uniqueness?", reviewed talk and abstract in *Current challenges and applications of comparative cognition (CompCog II)* Prague, 25–27 May 2011.

5. Joanna J. Bryson, "The Impact on Organisms of Transmitting Secondary Replicator Systems", with Sam P. Brown, reviewed abstract and talk, *The European Human Behaviour Association Conference*, Gießen, Germany 24–26 March 2011.

6. Joanna J. Bryson, "The Scientific Application of Agent-Based Modelling; The Impact on Organisms of Transmitting Secondary Replicator Systems", solicited talks for the workshop *Microbial Evolution: Modeling and Experimental Techniques*, e-Science Institute, University of Edinburgh, 11–13 January 2011.

7. Joanna J. Bryson, "Social Simulation and Explaining Religion", talk in the *The Explaining Religion Project (*EXREL*): How Do Religions Evolve?* symposium of *Religion, a Human Phenomena: The XXth Quinquennial World Congress of the International Association for the History of Religions*. Toronto, Canada, 17 August 2010.

8. Joanna J. Bryson, "Determinants of the Size of Social Species' Culture" talk and reviewed abstract, *Complexity and Nonlinear Phenomena in Biological Systems*, a one-day meeting organised by the Nonlinear and Complex Physics Group of the Institute of Physics (UK), Bath, 20 May 2010.

7.1.2 Invited

All below invited talks were by Bryson except where indicated.

1. "Cognition and Ecology: What Nature Tells Us About Autonomy", *Department of Informatics*, University of Edinburgh, 12 December 2011.

2. **"Understanding Cultural Variation in Social Behaviour: A Simulation-Based Approach to Anthropology"**, *Department of Computer Engineering*, Koç University, Istanbul, 2 December 2011.

3. "Memory, Action and Expectation — The Means and Extent to Which You and Your Society Construct Your Self", invited talk and network participant, *The Future of Memory*, The Memory Network, London, 29–30 October 2011.

4. "Space, Time and Information: Modelling the Ecology of Cultural Species", invited talk at *Modelling Ecological Environments and Agents: Understanding the Past and Predicting the Future* Hominid Dispersals Research Group, McGill and Montreal Universities, 24–26 October 2011.

5. "Cognition and Ecology: What Nature Tells Us About Autonomy", *Centre for Research in Cognitive Science (COGS)*, University of Sussex, 20 October 2011.

6. **Plenary** "Cognition and Ecology: What Nature Tells Us About Autonomy", *Autonomous Activity in the Real World: The Fifth EUCogII Members Conference*, Groeningen, 10 October 2011.

7. **"Evolution and information: The biology of social learning."** *Evolution@Bath&Bristol*, Bath, 26th September 2011.

8. **Plenary:** "Why We Share Information: The Evolution of Social Behaviour", *IJCAI 2011 Workshop on Agents Learning Interactively from Human Teachers (ALIHT)*, Barcelona, 16–17 July 2011.

9. "Sharing Information: The Adaptive Value of Social Behaviour", The Learning Algorithms and Systems Laboratory (LASA, Prof. Aude Billard), École Polytechnique Fédéarale de Lausanne (EPFL), Switzerland, 15 June 2011.

10. Benedikt Herrmann, **Plenary**: "Cooperation and cooperation enforcement among strangers is fundamental to the success" [sic], in *Current challenges and applications of comparative cognition (CompCog II)* Prague, 25–27 May 2011.

11. "Why Information Can Be Free: The Evolutionary Origins of Collective Intelligence", Human Dynamics Laboratory (Alex Pentland), The Media Laboratory, Massachusetts Institute of Technology (MIT), 17 May 2011.

12. "Modelling the evolution of social behaviour: Can sharing valuable information be adaptive?", Ecology and Evolution Group (Prof. Torben Dabelsteen), Department of Biology, University of Copenhagen, 13 May 2011.

13. "The Utility of Human Social Motivations: Biology, Individuality, and Culture", Department of Computing (Eduardo Alonso), City University London, 28 February 2011.

14. "Modelling social evolution: The role of space & time in the evolution of co-operation", Human Evolutionary Ecology Group (Prof. Ruth Mace), Department of Anthropology, University College of London, 25 January 2011.

15. "The Utility of Human Social Motivations: Biology, Individuality, and Culture", School of Computing and Mathematics (James Borg), Keele University, UK, 17 November 2010.

16. "The scientific application of agent-based modelling: From biology to anthropology", The Department of Software and Computer Science Education (Cyril Brom), Charles University, Prague, 19 November 2010.

17. "The scientific application of agent-based modelling: From biology to anthropology", invited talk at the *Interdisciplinary Workshop on Society, Culture and Language*, Plymouth, 11-13 November 2010.

18. "The Biological Basis of Human Social Motivations: Selective Pressure for Culture and Individuality", Faculty of Social Sciences, Carl von Ossietzky Universität Oldenburg, 31 May 2010.

19. Joanna J. Bryson and Will Lowe, "Agent-Based Modelling for Social Science: Modelling & Cultural Evolution", invited presentation to *The Second EXREL (Explaining Religion) Project Conference*, Centre for Anthropology and Mind, Oxford, UK; 17–19 August 2009.

20. Invited participant in *Integrating Cultures: Models, Simulations and Applications*, including presentation "A Computation-Enabled Biological Perspective on Cultural Variation", The Lorentz Center, Leiden, 6–9 April 2009.

21. **Plenary:** "Adaptive & Computational Explanations for The Pervasiveness of Social Learning and Altruistic Communication", at *Organisation, Cooperation and Emergence in Social Learning Agents*, a workshop at The Tenth European Conference on Artificial Life (ECAL '09), Budapest, 15 September 2009.

22. "Adaptive Trade-Offs Concerning Cognition and Culture", Oxford University, Department of Experimental Psychology, 16 June; also Oxford University, Centre for Anthropology & Mind, 17 June 2009.

23. "Understanding Human Social Motivations: Modeling the Evolution of Cognition and Culture", Central European University, Department of Philosophy, Budapest 2 June 2009.

7.1.3   Submitted

1. Daniel J. Taylor and Joanna J. Bryson, **"Rogers' Paradox: Resolved'**, reviewed abstract and talk, *The European Human Behaviour Association Conference*, Durham UK March 2012.

2. Simon T. Powers, Daniel J. Taylor and Joanna J. Bryson **"Punishment can promote defection in group-structured populations"**, reviewed abstract and talk, *The European Human Behaviour Association Conference*, Durham UK March 2012.

7.2   Publications

All of the below contain text pertaining explicitly to outcomes of the project, except for those in Section 7.2.4.

7.2.1   Available or In Press

1. Joanna J. Bryson "The Role of Stability in Cultural Evolution: Innovation and Conformity in Implicit Knowledge Discovery", in *Integrating Cultures: Formal Models and Agent-Based Simulations*, Virginia Dignum, Frank Dignum, Jacques Ferber, and Tiberiu Stratulat (eds), Springer, Berlin 2012 (in press).

2. Anil K. Seth, Tony J. Prescott, Joanna J. Bryson (eds.), *Modelling Natural Action Selection*, Cambridge University Press, 2012 (available since November 2011).

3. Gideon M. Gluckmann and Joanna J. Bryson, "An Agent-Based Model of the Effects of a Primate Social Structure on the Speed of Natural Selection", in *Evolutionary Computation and Multi-Agent Systems and Simulation (ECoMASS)*, Bill Rand and Forrest Stonedahl (eds.), Dublin 2011.

7.2.2   Submitted or In Revision

1. Harvey Whitehouse, Ken Kahn, Michael E. Hochberg and Joanna J. Bryson, "The role for simulations in theory construction for the social sciences: Case studies concerning Divergent Modes of Religiosity", submitted to *Method and Theory in the Study of Religion*.

2. Simon T. Powers, Daniel J. Taylor and Joanna J. Bryson, "Punishment can promote defection in group-structured populations", submitted to *The Journal of Theoretical Biology*.

3. Karolina Sylwester, James Mitchell and Joanna J. Bryson, "Punishment as aggression: Uses and consequences of costly punishment across populations", submitted to *Proceedings of the Royal Society: B—Biology*.

4. Daniel J. Taylor and Joanna J. Bryson, "Adaptation, The Fundamental Theorem of Natural Selection, and a Resolution to Rogers's Paradox", submitted to *Biological Philosophy*.

5. Joanna J. Bryson, Avri Bilovich, Ivana Čače, "Factors determining the extent of a species' reliance on socially-acquired behaviour", to be resubmitted to *Evolution of Human Behavior*.

6. Karolina Sylwester, James Mitchell and Joanna J. Bryson, "*Homo homini lupus*? Anti-social punishment revisited" to be resubmitted to *Human Nature*.

7.2.3   In Preparation

1. Karolina Sylwester, James Mitchell and Joanna J. Bryson, "Cooperative predispositions vary across populations and predict punishment behaviour".

2. Daniel J. Taylor, Simon T. Powers and Joanna J. Bryson, "Modularity in Cultural Evolution: The Utility of Limits to Group Size".

7.2.4   Related

Publications in this section were related to project goals, but could not be costed to this project nor contain AFOSR acknowledgement because they required human subjects.

- Jakub Gemrot, Cyril Brom, Joanna Bryson and Michal Bída, "How to compare usability of techniques for the specification of virtual agents behavior? An experimental pilot study with human subjects", in *Proceedings of the AAMAS 2011 Workshop on the uses of Agents for Education, Games and Simulations*, Taipei, M. Beer, C. Brom, V-W Soo and F. Dignum (eds.) May 2011.

- Michael S. Tremante "Behavioral Economics in On-line Environments" , undergraduate dissertation and technical report, University of Bath, 2011.

References

Abbink, K. and Sadrieh, A. (2009). The pleasure of being nasty. *Economics Letters*, 105(3):306–308.

Axelrod, R. and Hamilton, W. D. (1981). The evolution of cooperation. *Science*, 211:1390–1396.

Azevedo, R., Lohaus, R., Srinivasan, S., Dang, K., and Burch, C. (2006). Sexual reproduction selects for robustness and negative epistasis in artificial gene networks. *Nature*, 440(7080):87.

Baker, M. and Marler, P. (1980). *Evolution of social behavior — hypotheses and empirical tests : report of the Dahlem Workshop on Evolution of Social Behavior*. Verlag Chemie, Weinheim, Berlin.

Baldwin, J. M. (1896). A new factor in evolution. *The American Naturalist*, 30:441–451.

Bowles, S. (2009). Did warfare among ancestral hunter-gatherers affect the evolution of human social behaviors? *Science*, 324(5932):1293–1298.

Boyd, R. and Richerson, P. (1996). Why culture is common, but cultural evolution is rare. In *Proceedings of the British Academy*, volume 88, pages 77–94. Oxford University Press.

Boyd, R. and Richerson, P. J. (1995). Why does culture increase human adaptability? *Ethology and Sociobiology*, 16(2):125–143.

Boyd, R., Richerson, P. J., and Henrich, J. (2011). The cultural niche: Why social learning is essential for human adaptation. *Proceedings of the National Academy of Sciences*, 108(Supplement 2):10918.

Bryson, J. J. (2008). The role of modularity in stabilizing cultural evolution: Conformity and inno-vation in an agent-based model. In Davis, A. and Ludwig, J., editors, *Proceedings of the AAAI Fall Symposium on Adaptive Agents in Cultural Contexts (AACC '08)*, pages 8–17, Arlington, VA. AAAI Press.

Bryson, J. J. (2009). Representations underlying social learning and cultural evolution. *Interaction Studies*, 10(1):77–100.

Bryson, J. J., Ando, Y., and Lehmann, H. (2007). Agent-based models as scientific methodology: A case study analysing primate social behaviour. *Philosophical Transactions of the Royal Society, B — Biology*, 362(1485):1685–1698.

Bryson, J. J., Ando, Y., and Lehmann, H. (2012). Agent-based models as scientific methodology: A case study analyzing the DomWorld theory of primate social structure and female dominance. In Seth, A. K., Prescott, T. J., and Bryson, J. J., editors, *Modelling Natural Action Selection*, pages 427–453. Cambridge University Press.

Di Fiore, A. and Rendall, D. (1994). Evolution of social organization: A reappraisal for primates by using phylogenetic methods. *Proceedings of the National Academy of Sciences*, 91:9941–9945.

Enquist, M., Eriksson, K., and Ghirlanda, S. (2007). Critical social learning: A solution to Rogers's paradox of nonadaptive culture. *American Anthropologist*, 109(4):727–734.

Fehr, E. and Gächter, S. (2000). Cooperation and punishment in public goods experiments. *The American Economic Review*, 90(4):980–994.

Fehr, E. and Gächter, S. (2002). Altruistic punishment in humans. *Nature*, 415(6868):137–140.

Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin*, 101(2):171.

Gintis, H., Bowles, S., Boyd, R., and Fehr, E. (2003). Explaining altruistic behavior in humans. *Evolution and Human Behavior*, 24(3):153–172.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., and McElreath, R. (2001). Cooperation, reciprocity and punishment in fifteen small-scale societies. *American Economic Review*, 91(2):73–78.

Henrich, J., Heine, S. J., and Norenzayan, A. (2010). Most people are not WEIRD. *Nature*, 466(7302):29.

Henrich, J. and McElreath, R. (2003). The evolution of cultural evolution. *Evolutionary Anthropology: Issues, News, and Reviews*, 12(3):123–135.

Herrmann, B., Thöni, C., and Gächter, S. (2008a). Antisocial punishment across societies. *Science*, 319(5868):1362–1367.

Herrmann, B., Thöni, C., and Gächter, S. (2008b). Supporting online material for antisocial punishment across societies. *Science*, 319(5868). http://www.sciencemag.org/cgi/content/full/319/5868/1362/DC1.

Hinton, G. E. and Nowlan, S. J. (1987). How learning can guide evolution. *Complex Systems*, 1:495–502.

Hobbes, T. (1651). *Leviathan*. Andrew Crooke, London.

Jensen, K. (2010). Punishment and spite, the dark side of cooperation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1553):2635–2650.

Kaufmann, D., Kraay, A., and Mastruzzi, M. (2004). Governance matters III: Governance indicators for 1996, 1998, 2000, and 2002. *The World Bank Economic Review*, 18(2):253–287.

Kokko, H. (2007). *Modelling for Field Biologists and Other Interesting People*. Cambridge University Press.

Laland, K. (2004). Social learning strategies. *Learning & Behavior*, 32(1):4.

Lamba, S. and Mace, R. (2011). Demography and ecology drive variation in cooperation across human populations. *Proceedings of the National Academy of Sciences*, 108(35):14426–14430.

Laver, M. J. (2005). Policy and the dynamics of political competition. *American Political Science Review*, 99(2):263–281.

Lehmann, L., Feldman, M. W., and Kaeuffer, R. (2010). Cumulative cultural dynamics and the coevolution of cultural innovation and transmission: an ESS model for panmictic and structured populations. *Journal of Evolutionary Biology*, 23(11):2356–2369.

MacLean, R. C., Fuentes-Hernandez, A., Greig, D., Hurst, L. D., and Gudelj, I. (2010). A mixture of "cheats" and "co-operators" can enable maximal group benefit. *PLoS Biol*, 8(9):e1000486.

Perelman, P., Johnson, W. E., Roos, C., Seunez, H. N., Horvath, J. E., Moreira, M. A. M., Kessing, B., Pontius, J., Roelke, M., Rumpler, Y., Schneider, M. P. C., Silva, A., O'Brien, S. J., and Pecon-Slattery, J. (2011). A molecular phylogeny of living primates. *PLoS Genetmal*, 7(3):e1001342.

Powell, A., Shennan, S., and Thomas, M. G. (2009). Late Pleistocene demography and the appearance of modern human behavior. *Science*, 324(5932):1298–1301.

Powers, S. T., Penn, A. S., and Watson, R. A. (2011). The concurrent evolution of cooperation and the population structures that support it. *Evolution*, 65(6):1527–1543.

University of Bath, Department of Computer Science   *J J Bryson, PI*

Preuschoft, S. and van Schaik, C. P. (2000). Dominance and communication: Conflict management in various social settings. In Aureli, F. and de Waal, F. B. M., editors, *Natural Conflict Resolution*, chapter 6, pages 77–105. University of California Press.

Purvis, A. (1995). A composite estimate of primate phylogeny. *Philosophical Transactions: Biological Sciences*, 348(1326):405–421.

Rendell, L., Fogarty, L., and Laland, K. N. (2009). Rogers' paradox recast and resolved: Population structure and the evolution of social learning strategies. *Evolution*, 64(2):534–548.

Richerson, P. J. and Boyd, R. (2005). *Not By Genes Alone: How Culture Transformed Human Evolution*. University Of Chicago Press.

Rogers, A. (1988). Does biology constrain culture? *American Anthropologist*, 90(4):819–831.

Sperber, D. and Hirschfeld, L. A. (2004). The cognitive foundations of cultural stability and diversity. *Trends in Cognitive Sciences*, 8(1):40–46.

Szathmáry, E. (2011). To group or not to group? *Science*, 334(6063):1648–1649.

Čače, I. and Bryson, J. J. (2007). Agent based modelling of communication costs: Why information can be free. In Lyon, C., Nehaniv, C. L., and Cangelosi, A., editors, *Emergence and Evolution of Linguistic Communication*, pages 305–322. Springer, London.

West, S. A., Griffin, A. S., and Gardner, A. (2007). Evolutionary explanations for cooperation. *Current Biology*, 17:R661–R672,.