Wrath of God: Religious Primes and Punishment

RYAN MCKAY[1,2,3,4], CHARLES EFFERSON[1,2], HARVEY WHITEHOUSE[3] & ERNST FEHR[1,2]

[1] Institute for Empirical Research in Economics, University of Zürich, Switzerland
[2] Laboratory for Social and Neural Systems Research, University of Zürich, Switzerland
[3] Centre for Anthropology and Mind, University of Oxford, UK
[4] Department of Psychology, Royal Holloway, University of London, UK

Please address correspondence to:
Dr Ryan McKay
Department of Psychology
Royal Holloway, University of London
Egham, Surrey TW20 0EX
United Kingdom
Tel: +44 (0) 1784 41 4993
Email: ryantmckay@mac.com

Summary

Recent evidence indicates that priming participants with religious concepts promotes prosocial sharing behaviour. In the present study we investigated whether religious priming also promotes the costly punishment of unfair behaviour. 304 participants played a punishment game. Before the punishment stage began, participants were subliminally primed with religion primes, secular punishment primes, or control primes. We found that religious primes strongly increased the costly punishment of unfair behaviours for a subset of our participants – those who had previously donated to a religious organization. We discuss two proximate mechanisms potentially underpinning this effect. The first is a 'supernatural watcher' mechanism, whereby religious participants punish unfair behaviours when primed because they sense that not doing so will anger or disappoint an observing supernatural agent. The second is a 'behavioural priming' mechanism, whereby religious primes activate cultural norms pertaining to fairness and its enforcement and occasion behaviour consistent with those norms. We conclude that our results are consistent with dual inheritance proposals about religion and cooperation, whereby religions harness the byproducts of genetically inherited cognitive mechanisms in ways that enhance the survival prospects of their adherents.

Key words: Religion; Supernatural Agency; Subliminal Priming; Fairness; Punishment; Cooperation.

Introduction

The LORD is a jealous God, filled with vengeance and wrath… (Nahum 1:2)

Religion carries formidable epistemic, metabolic and material costs [1-3]. Religious believers must maintain and compartmentalize beliefs that are extravagantly at variance with intuitive conceptions of reality. Religious rituals, moreover, are often physically taxing and painful, and frequently require the sacrifice of precious resources. Given such costs, some evolutionary theorists argue that religion must provide, or in the ancestral past must have provided, countervailing adaptive benefits [e.g., 1-7; cf. 8-9]. Perhaps the most influential of such proposals is that religion is a cultural variant that confers a selective advantage at the group level by virtue of the fact that it secures and promotes cooperative behaviour within the group [6-7]. This proposal arguably solves not one but two thorny evolutionary puzzles: The puzzle of religion and the puzzle of human cooperation.

The nature and extent of human cooperation is unique in the animal kingdom [10-11]. Human societies are based on large-scale cooperation between genetically unrelated individuals. Cooperation is frequent in non-repeated interactions, even when reputational gains are small or absent. Cooperation and other prosocial behaviours will in many situations be sustained by preferences for *fairness*, or by a cultural norm of fairness [10]. Such preferences are evidenced by the behaviour of participants in anonymous, one-shot economic games, many of who nominate fair outcomes even when such outcomes are disadvantageous with respect to their material self-interest [10,12]. Humans, moreover, reward others who behave fairly and impose sanctions on those who fail to do so [10].

One potential means of implementing fairness norms is via culturally postulated supernatural agents [9], in particular "full-access strategic agents" such as omnipotent, omniscient, moralizing gods [8,13]. Individuals who believe that behavioural norms are policed by an all-knowing supernatural agent with the power and inclination to inflict terrible retribution for norm violations will have a strong incentive to comply with those norms. Some authors, therefore (e.g., 7,14-16], have suggested that belief in supernatural punishment confers a selective advantage by promoting prosocial behaviour.

Recent evidence from studies employing priming paradigms is consistent with this proposal. In a seminal study, Shariff and Norenzayan [17] used a scrambled sentence task to prime religious concepts, and found that participants primed in this fashion gave significantly more money in a subsequent (anonymous, one-shot) dictator game than did control participants. Similar results have been found in other recent priming studies. For example, relative to control participants, participants primed with religious or supernatural concepts have been found to cheat less [18-19], to collect more charity pamphlets [20], and to be more likely to cooperate in a prisoner's dilemma game [21].

Religious priming appears to promote prosocial behaviour – but does it also promote the costly punishment of unfair behaviour? In the present study we sought to investigate this issue. In order to minimise demand characteristics [see 22], we decided to present primes subliminally. Our research questions were threefold: 1) Would participants primed with the concepts of *religion* and/or *punishment* punish more in a punishment game? 2) Would such primes influence punishment of unfair behaviours only, or punishment of both unfair *and* fair behaviours? 3) Would any effects of religious primes be limited to participants with religious commitments?

Methods

Participants and General Procedure
The sample comprised 304 participants (140 females, 164 males; Mean age ± *SD* = 21.9 years ± 3.7), most of who were students at the University of Zürich or the Swiss Federal Institute of Technology in Zürich. The breakdown of religious affiliations broadly mirrored that of Zürich society in general[1] and was as follows: approximately 30% Protestant, 28% Catholic and 42% other affiliations/no affiliation. Recruitment was conducted using the Online Recruitment System for Economic Experiments [ORSEE; 23].

The experimental procedure was as follows. Participants were randomly assigned the role of either Player A or Player B and played a two-stage punishment game with a player of the opposite type. Between the two stages of the game participants underwent a subliminal priming episode, and after the completion of the second stage they undertook a systematic test of prime visibility. Finally, participants filled out two questionnaires – one to collect demographic information (age, gender etc.) and one to collect information about religious affiliation, beliefs and practices. Upon completion, participants were paid a show up fee of 10 Swiss Francs (CHF) plus their earnings from the experiment.

Punishment Game
We measured punishment using a two-player second party punishment game [24]. This game had a two-stage structure. In the first stage, Player A chose an allocation of proposed payoffs to herself and Player B. Allocations were presented in points (1 point = 0.28 CHF). Two options were presented on the computer screen for Player A to choose between: a fair option (150,150) and an unfair option (590,60). In each option the values on the left and right indicated the shares of Players A and B respectively.

In the second stage of the game Player B was informed of the two options that were available to Player A in the first stage, but did not learn the specific choice that Player A made. Instead, we used the 'strategy method' in order to maximize the amount of

---

statistical data gathered: Player B was, for each option, given the opportunity to spend points out of her allocation share in order to reduce Player A's payoff in that case - i.e., to punish Player A. In each case the choice was binding, provided that the relevant option was actually chosen by Player A. Previous work has shown that participants' qualitative behavioural patterns are unaffected by the use of this method as opposed to the 'direct-response' method, in which Player B learns the specific choice made by Player A and then chooses a response [25-26]. Player B could spend a maximum of 50 points (and minimum of 0) to punish Player A, i.e., Players B entered a number between 0 and 50 for each of the two options. A 1:3 punishment ratio was employed, such that each point spent by Player B reduced the payoff of Player A by three points - provided that the relevant choice was actually made by Player A. Hence, if Player A chose the allocation (x_A, x_B) and Player B punished her with $0 <= p <= 50$ points for that choice, Player A's payoff was $x_A - 3p$ and Player B's payoff was $x_B - p$.

Priming Episode and Visibility Check

The priming episode followed the first stage of the punishment game. There were four between-subjects priming treatments:

1) Religion (primes: *divine, holy, pious, religious*)
2) Punishment (primes: *revenge, punish, penalty, retribution*)
3) Religion-Punishment (primes: *divine, revenge, pious, punish*)
4) Control (primes: *northeast, acoustic, tractor, carton*)[2]

The priming episode comprised 20 priming trials. The sequence of events for each trial was as follows (see Figure 1): fixation point for 500ms, forward mask for 500ms, prime for 40ms then backward mask for 500ms. As soon as the priming episode concluded the second stage of the punishment game began, and once this stage was complete participants underwent a systematic test for prime visibility. We excluded any participants who performed significantly above chance on this test.

Insert Figure 1 about here

Religion Questionnaire

In addition to requesting religious affiliation, our religion questionnaire included a series of items answered on a five-point Likert scale (1=Strongly Disagree; 5=Strongly Agree; see Table 1 for a list of these items) followed by a single YES/NO item: "In the

---

[2] Primes were presented in German; these are English translations. See Table S1 in the supplementary material available online for the German words actually presented.

past year, did you donate to a religious organization?"

Insert Table 1 about here

Results

We ran separate analyses for two dependent variables:

      1. Punishment of the unfair choice (590,60); See Table 2.
      2. Punishment of the fair choice (150,150); See Table 3.

In each case we analysed the actual strength (amount) of punishment rather than simply whether punishment occurred or not at any strength.

Insert Tables 2 and 3 about here

There were no significant main effects of our experimental primes on punishment of the unfair (Table 2, Model 1) or fair (Table 3, Model 1) choices, although the effect of *Punishment* primes was marginally significant (and positive) for punishment of the unfair choice (see coefficient for 'Punishment Prime' in Table 2, Model 1). For punishment of the unfair choice, however, there was a highly significant interaction between religious donations and the *Religion* priming treatment (see coefficient for 'Religious Donations x Rel' in Table 2, Model 2). Neither religious donations nor religious priming had effects in isolation (see respective coefficients for 'Religious Donations' and 'Religion Prime' in Table 2, Model 2), but when both were present they jointly increased punishment of the unfair choice by .84 of a standard deviation, $p = 0.001$[3]. Apart from a significant main effect of gender for punishment of the unfair choice (see coefficient for 'Female' in Table 2, Model 2: in line with previous research, females were less punitive; [see, e.g., 27]), no other effects were significant for either type of punishment.

---

[3] Note that the coefficients in our models are in natural units. When we re-analyse with normalized continuous variables the coefficient for 'Religious Donations x Rel' is .84.

Discussion

Humans are subject to strong cultural norms of fairness. A substantial proportion of participants in anonymous, one-shot economic games nominate fair outcomes even when such outcomes are to their material disadvantage [10,12]. Humans, moreover, incur costs to reward others who behave fairly and to impose sanctions on those who behave unfairly [11]. Recent studies indicate that priming participants with religious concepts promotes prosocial behaviour [e.g., 17,19]. Our aim in the present experiment was to investigate whether religious priming would also promote the costly punishment of unfair behaviour.

Across all participants, our results indicate a negative answer to this question: There was no main effect of religious primes, whether alone or in combination with punishment primes, on punishment behaviour. The only suggestion of a main effect for priming treatment was for punishment primes alone, which (perhaps unsurprisingly) tended to increase punishment of unfair choices. However, religious primes did *strongly* increase the costly punishment of unfair behaviour for a subset of our participants – those who had previously donated to a religious organization.

How are we to account for these results? In line with Shariff and Norenzayan [17], we consider two possible explanations. The first is that religious primes activate the notion that one's behaviour is observed by a supernatural agent. In this case primed participants punish unfair behaviours because they sense that not doing so will damage their standing in the eyes of a supernatural agent. Recent studies suggest that even very subtle cues that one is being watched, such as stylized eyespots on a computer screen, can affect giving behaviour [e.g., 28-29; cf. 30]. Some authors have suggested that such cues match the input conditions for evolved mental mechanisms that detect when one's behaviour is observed [28]. Religious primes might likewise function as input for these mechanisms [17].

The second possibility is a behavioural priming explanation, whereby religious primes activate cultural norms pertaining to fairness and its enforcement and occasion behaviour consistent with those norms. This explanation is consistent with evidence that the activation of conceptual representations increases the likelihood of behaviours consistent with those representations [e.g., 31]. Thus, much as participants walk more slowly down a length of corridor when the concept "elderly" is primed [31], priming words that are semantically associated with fairness may lead participants to punish unfair behaviours simply by virtue of that semantic connection [22]. Bargh et al. [32] found that participants primed with cooperation-related words (e.g., *fair, share*) were less selfish in a subsequent resource-management game, and Shariff and Norenzayan [17] found that priming with secular-moral words (e.g., *court, contract*) had a similar effect to that of religious primes on fair allocations in a subsequent dictator game.

Although they acknowledge that the two mechanisms above need not be mutually

exclusive, Shariff and Norenzayan [17] favour the 'supernatural watcher' account (cf. 22). Norenzayan et al. [33] argue that behavioural priming effects are "typically impervious to prior explicit beliefs or attitudes" (p. 532). If this is true, then one would not expect the effects of religious primes to be mediated by individual religiosity if those effects are attributable to behavioural priming. Norenzayan et al. [33] also describe a recent series of studies which found that religious primes caused an increase in public self-awareness, which is "characterized by attentiveness to those features of one's self that are presented to others" [34, p. 366] and so directly linked to sensitivity about being observed.

Regarding the former point, it seems to us that the effect of activating a certain set of cultural norms might well be stronger for those who have internalized those norms. With respect to our experiment, it seems plausible that the behavioural norms of religious institutions are more strongly represented in the minds of individuals who financially support those institutions, and thus more susceptible to activation by relevant primes. Nevertheless, we agree with Norenzayan et al. [33] that multiple psychological mechanisms may be operative and even mutually reinforcing. If an individual believes that in order to avoid punishment herself she needs both to adhere to and to uphold cultural norms of fairness, then religious primes may affect punishment behaviours both by evoking a sense of being observed and by directly activating the relevant norms. Future work might profitably investigate these possibilities.

The accounts we have considered above pertain to proximate explanation. In terms of ultimate evolutionary explanation, our results are consistent with dual inheritance proposals about religion and cooperation.[4] A number of authors [e.g., 36-39] have suggested that the human proclivity for acquiring and transmitting supernatural agent concepts is an incidental byproduct of cognitive mechanisms genetically adapted for other purposes. Others [e.g., 33,40-43] have argued that religions are cultural systems that exploit such byproducts to adaptive effect. If, as our results indicate, the activation of supernatural agent concepts promotes the enforcement of cultural norms of fairness, and if such norms sustain cooperative behaviours within the group, then religions that harness such concepts will enhance the survival prospects of their vectors, thereby contributing to their own survival.

---

[4] Recently, Henrich, Heine and Norenzayan [35] documented evidence that university students, particularly in Western societies, are frequent outliers on many psychological measures. Given that most of our participants were university students in Zürich, a note of caution about the generalisabity of our findings is in order.

Acknowledgements

References

1. Bulbulia, J. (2004). The cognitive and evolutionary psychology of religion. *Biol. Philos.,* **19**, 655-686.
2. Murray, M. J., & Moore, L. (2009). Costly signaling and the origin of religion. *Journal of Cognition and Culture,* **9**, 225-245.
3. Sosis, R. (2004). The adaptive value of religious ritual. *American Scientist,* **92**, 166-172.
4. Alcorta, C. & Sosis, R. (2005). Ritual, emotion, and sacred symbols: The evolution of religion as an adaptive complex. *Hum. Nat.,* **16**, 323-59.
5. Johnson, D. D. P. & Bering, J. M. (2006). Hand of God, mind of man: Punishment and cognition in the evolution of cooperation. *Evolutionary Psychology,* **4**, 219-33.
6. Norenzayan, A. & Shariff, A. F. (2008). The origin and evolution of religious prosociality. *Science,* **322**, 58-62.
7. Wilson, D. S. (2002). *Darwin's cathedral: Evolution, religion and the nature of society.* Chicago: University of Chicago Press.
8. Boyer, P. (2001). *Religion Explained: The Evolutionary Origins of Religious Thought.* New York: Basic Books.
9. McCauley, R. N. & Lawson, E. T. (2002). *Bringing ritual to mind.* Cambridge University Press.
10. Fehr, E. & Fischbacher, U. (2003). The nature of human altruism. *Nature,* **425**, 785-791.
11. Fehr, E. & Gächter, S. (2002). Altruistic punishment in humans. *Nature,* **415**, 137-140.
12. Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and co-operation. *Q. J. Econ.,* **114**, 817-868.
13. Bering, J. M. (2006). The folk psychology of souls. *Behav. Brain Sci.,* **29**, 453-498.
14. Henrich, J., Ensminger, J., McElreath, R., Barr, A., Barrett, C., Bolyanatz, A., Cardenas, J. C., Gurven, M., Gwako, E., Henrich, N. et al. (2010). Markets, religion, community size, and the evolution of fairness and punishment. *Science,* **327**, 1480-1484.
15. Johnson, D. D. P. & Krüger, O. (2004). The good of wrath: Supernatural punishment and the evolution of cooperation. *Political Theology,* **5**, 159-76.
16. Shariff, A., Norenzayan, A. & Henrich, J. (2010). The birth of high gods. In M. Schaller, A. Norenzayan, S. J., Heine, T. Yamagishi, & T. Kameda (Eds.), *Evolution, Culture, and the Human Mind.* (pp. 119-136). Psychology Press - Taylor & Francis.
17. Shariff, A. F. & Norenzayan, A. (2007). God is watching you: Priming God concepts increases prosocial behavior in an anonymous economic game. *Psychol. Sci.,* **18**, 803-809.
18. Bering, J. M., McLeod, K. A. & Shackelford, T. K. (2005). Reasoning about dead agents reveals possible adaptive trends. *Hum. Nat.,* **16,** 360-81.
19. Randolph-Seng, B. & Nielsen, M. E. (2007). Honesty: One effect of primed religious representations. *The Int. J. Psychol. Relig.,* **17**, 303-315.

20. Pichon, I., Boccato, G. & Saroglou, V. (2007). Nonconscious influences of religion on prosociality: A priming study. *Eur. J. Soc. Psychol.*, **37**, 1032–1045.

21. Ahmed, A. M. & Salas, O. (2008). In the back of your mind: Subliminal influences of religious concepts on prosocial behavior. Working paper.

22. Randolph-Seng, B. & Nielsen, M. E. (2008). Is God really watching you? A response to Shariff and Norenzayan (2007). *The Int. J. Psychol. Relig.*, **18**, 119-122.

23. Greiner, B. (2004). The Online Recruitment System ORSEE 2.0: A guide for the organization of experiments in economics. *Working Paper Series in Economics 10, University of Cologne, Department of Economics.*

24. Leibbrandt, A. & López-Pérez, R. (2008). The envious punisher: Understanding third and second party punishment with simple games. *IEW – Working Papers,* iewwp373, Institute for Empirical Research in Economics - IEW.

25. Brandts, J. and G. Charness (2009), The strategy versus the direct-response method: A survey of experimental comparisons, Discussion Paper.

26. Falk, A., Fehr, E. & Fischbacher, U. (2005). Driving forces behind informal sanctions. *Econometrica*, **7**, 2017-30.

27. Eckel, C. C. & Grossman, P. J. (2001). Chivalry and solidarity in ultimatum games. *Economic Inquiry,* **39**, 171-188.

28. Haley, K. J. & Fessler, D. M. T. (2005). Nobody's watching? Subtle cues can affect generosity in an anonymous economic game. *Evol. Hum. Behav.*, **26**, 245-256.

29. Rigdon, M., Ishii, K., Watabe, M. & Kitayama, S. (2009). Minimal social cues in the dictator game. *Journal of Economic Psychology,* **30**, 358 – 367.

30. Fehr, E. & Schneider, F. (2010). Eyes are on us, but nobody cares: Are eye cues relevant for strong reciprocity? *Proc. R. Soc. B,* **277**, 1315-1323.

31. Bargh, J. A., Chen, M. & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *J. Pers. Soc. Psychol.*, **71**, 230–244.

32. Bargh, J. A., Gollwitzer, P. M., Lee-Chai, A., Barndollar, K. & Troetschel, R. (2001). Automating the will: Nonconscious activation and pursuit of behavioral goals. *J. Pers. Soc. Psychol.*, **81**, 1014–27.

33. Norenzayan, A., Shariff, A. F. & Gervais, W. M. (2009). The evolution of religious misbelief. *Behav. Brain Sci.*, **32**, 531-2.

34. Govern, J. M. & Marsch, L. A. (2001). Development and validation of the Situational Self-Awareness Scale. *Conscious. Cogn.*, **10**, 366-78.

35. Henrich, J., Heine, S. J. & Norenzayan, A. (2010). The weirdest people in the world? *Behav. Brain Sci.*, **33**, 61-83.

36. Barrett, J. L. (2000). Exploring the natural foundations of religion. *Trends Cogn. Sci.*, **4**, 29-34.

37. Bloom, P. (2004). *Descartes' baby: How child development explains what makes us human*. Arrow Books.

38. Bloom, P. (2007). Religion is natural. *Developmental Science,* **10**, 147-51.

39. Guthrie, S. E. (1993). *Faces in the clouds: A new theory of religion*. Oxford University Press.

40. Pyysiäinen, I. & Hauser, M. (2010). The origins of religion: Evolved adaptation or

by-product? *Trends Cogn. Sci.,* **14**, 104-9.

41. Richerson, P. & Boyd, R. (2005). *Not by genes alone: How culture transformed human evolution.* University of Chicago Press.

42. Talmont-Kaminski, K. (2009). Effective untestability and bounded rationality help in seeing religion as adaptive misbelief. *Behav. Brain Sci.,* **32**, 536-7.

43. Wilson, D. S. & Lynn, S. J. (2009). Adaptive misbeliefs are pervasive, but the case for positive illusions is weak. *Behav. Brain Sci.,* **32**, 539-40.

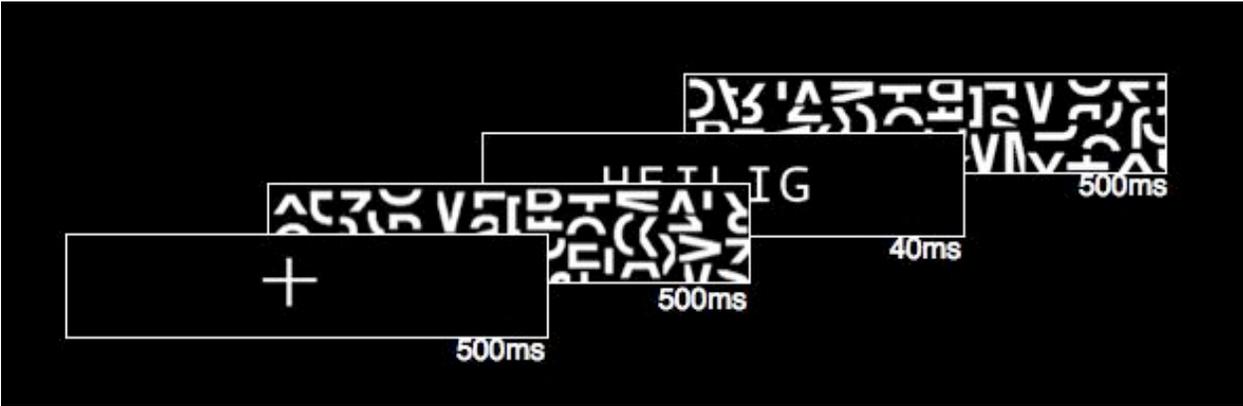Figure 1: Sequence of events on each priming trial

Table 1
Likert Items from Religion Questionnaire

---

1. I often think about God
2. I often attend religious services (apart from weddings, funerals & christenings)
3. I often pray outside of religious services
4. I often read or study religious texts outside of religious services
5. I believe in God
6. I believe in life after death
7. I believe God knows everything we do or think
8. I believe in heaven
9. I believe in hell
10. I believe God will punish sinners
11. I believe God will reward believers

Table 2

Models 1 and 2 for punishment of the unfair choice. Predictor variables include age and a dummy variable denoting female gender; a composite of the 11 Likert items from the religion questionnaire, representing the average of responses to these items; a dummy variable for religious donations; and dummies for the three experimental priming treatments. Each of the priming treatment dummy variables is also interacted with the Likert composite and with the religious donations dummy. An asterisk indicates significance at the 0.05 level, and two asterisks indicate significance at the 0.01 level.

Variable	Model 1			Model 2		
	Estimate	*SE*	*p* value	Estimate	*SE*	*p* value
Intercept	6.966	1.966	< 0.001 **	12.017	7.663	0.118
Age	____			0.078	0.261	0.766
Female	____			-5.248	2.023	0.010 *
Likert Composite	____			-1.093	2.112	0.605
Religious Donations	____			-7.054	5.264	0.182
Religion Prime	-0.170	2.844	0.952	-4.192	7.251	0.564
Punishment Prime	5.275	2.793	0.060	-5.508	6.808	0.419
Religion-Punishment Prime	1.960	2.844	0.492	-5.051	7.367	0.494
Likert Composite x Rel	____			0.131	3.102	0.966
Likert Composite x Pun	____			3.942	2.784	0.158
Likert Composite x Rel-Pun	____			2.566	2.993	0.392
Religious Donations x Rel	____			29.394	8.929	0.001 **
Religious Donations x Pun	____			3.377	7.587	0.657
Religious Donations x Rel-Pun	____			0.691	7.509	0.927

Table 3

Models 1 and 2 for punishment of the fair choice. Predictor variables are as per Table 2.

Variable	Model 1			Model 2		
	Estimate	*SE*	*p* value	Estimate	*SE*	*p* value
Intercept	1.610	1.314	0.222	2.515	5.278	0.634
Age	_____			-0.094	0.180	0.600
Female	_____			2.309	1.393	0.099
Likert Composite	_____			0.153	1.454	0.917
Religious Donations	_____			-1.643	3.625	0.651
Religion Prime	0.760	1.901	0.690	-6.059	4.993	0.226
Punishment Prime	3.079	1.867	0.100	-0.948	4.689	0.840
Religion-Punishment Prime	2.131	1.901	0.264	3.129	5.073	0.538
Likert Composite x Rel	_____			3.020	2.137	0.159
Likert Composite x Pun	_____			1.489	1.917	0.438
Likert Composite x Rel-Pun	_____			-0.165	2.061	0.936
Religious Donations x Rel	_____			-0.391	6.149	0.949
Religious Donations x Pun	_____			2.031	5.225	0.698
Religious Donations x Rel-Pun	_____			-2.190	5.171	0.672