

## Original Article

Does religious cognition really down-regulate hand grip endurance in men? A failure to replicate<sup>☆</sup>Liana S.E. Hone, Michael E. McCullough<sup>\*</sup>

Department of Psychology, University of Miami

## ARTICLE INFO

## Article history:

Initial receipt 7 April 2014

Final revision received 27 August 2014

## Keywords:

Replication

Religious prime

Hand grip endurance

Sexual selection

Reproductive Religiosity Model

## ABSTRACT

We sought to perform an exact replication of a previously published experiment that indicated that religious cognition (manipulated via an implicit religious prime) reduced hand-grip endurance in men but not in women. We randomly assigned 168 female and 159 male undergraduate students to either a task in which they completed scrambled sentences interspersed with words that had religious meanings or a comparable task with words that had no consistent meaning. We did not find an interaction between condition and sex: Men who received the religious prime did not perform any worse on the hand-grip endurance task than did their peers who completed the control task. We did, however, find a sex difference in hand-grip endurance (favoring men), but this sex difference was not found in the original experiment. We explored potential explanations for this failure to replicate, including the possibility that different experimenters (who were ignorant of participants' condition assignments) produced different results by affecting participants' motivation in different ways. These latter analyses revealed that some experimenters produced larger sex differences than others did.

© 2015 Elsevier Inc. All rights reserved.

## 1. Introduction

Some social scientists have recently argued that people commonly put their religious beliefs and religious group memberships to strategic use in the contemporary United States, and perhaps elsewhere, to support restricted reproductive strategies (Putnam & Campbell, 2012; Weeden, Cohen, & Kenrick, 2008). Restricted reproductive strategists promulgate high-fertility, committed, monogamous relationships, which are characterized by high parental investment and mate fidelity, in lieu of unrestricted, promiscuous relationships (Weeden et al., 2008). Evolutionary psychologists have argued that religious groups might facilitate monogamy and parental investment by providing reproductive support to families and by promulgating social norms that encourage fidelity (Weeden & Kurzban, 2013; Weeden et al., 2008). If support of restricted reproductive strategies is at the heart of variations in religiosity in contemporary society, we would expect reproductive morals to be more closely linked to religiosity than to other types of morality (e.g., cooperative morality), and this indeed appears to be the case—not only in the United States, but in virtually every other world region as well (Weeden & Kurzban, 2013).

If restricted reproductive religious environments do, in fact, deter promiscuous sexual strategists (Weeden et al., 2008), then exposure to cues associated with those environments might also be expected to down-regulate the expression of traits associated with promiscuous sexual strategies, and to up-regulate the expression of traits associated

with the pursuit of monogamous strategies (McCullough, Carter, DeWall, & Corrales, 2012). For men, examples of behaviors associated with the pursuit of unrestricted sexual strategies might include outlays of physical strength, risky demonstrations, and unwillingness to delay gratification (McCullough et al., 2012). McCullough et al. (2012) sought to test this idea by experimentally manipulating religious cognition and then observing its effects on two sex-differentiated behavioral traits—willingness to delay gratification and outlays of physical endurance (Archer, 2009; Daly & Wilson, 2005; Hawkes, 1991; Kirby & Maraković, 1996; Little & Johnson, 1986; Pawlowski, Atwal, & Dunbar, 2008; Shih, 2007; Silverman, 2003; Wilson & Daly, 2004).

In their first experiment (N = 180), McCullough et al. (2012) assigned participants to write an essay on their religion and God (religious prime condition), on their country and culture (secular condition), or on their household items (control condition). Upon completing the essay, participants were asked to complete the Monetary Choice Questionnaire (Kirby & Maraković, 1996) – a gauge of the degree to which participants hypothetically prefer small amounts of money sooner compared to larger amounts later – the results of which were used to calculate the rate at which participants discount future rewards. McCullough et al. (2012) found a significant sex by condition interaction indicating that men in the religious condition reported more willingness to delay gratification than did men in the secular and control conditions combined (the same was not true of women). In a second experiment (N = 171), McCullough et al. (2012) conceptually replicated their *Experiment 1* findings using a new religious priming method. Participants either read an essay ostensibly providing evidence for (religious prime condition) or against (non-religious prime condition) the existence of an afterlife, followed by the Monetary Choice

<sup>☆</sup> This study was funded by the John Templeton Foundation.

<sup>\*</sup> Corresponding author. PO Box 248185 Coral Gables, FL 33124-0751.

E-mail address: [mikem@miami.edu](mailto:mikem@miami.edu) (M.E. McCullough).

Questionnaire for which the rewards were merely hypothetical. McCullough et al. (2012) again found a sex by condition interaction indicating that men in the religious condition were less impulsive than men in the non-religious condition (the same was not true of women).

In their third and final experiment ( $N = 160$ ), McCullough et al. (2012) used a scrambled sentence task which comprised religious words (religious prime condition) or neutral words (control condition) to implicitly manipulate religious cognition. They then asked participants to maintain 70% of their maximum voluntary contraction score (which was estimated prior to the experimental manipulation) on a hand dynamometer for as long as possible. McCullough et al. (2012) found a sex by condition interaction indicating that men in the religious prime condition displayed a reduced maximum endurance time in comparison to men in the control condition (the same was not true of women). McCullough et al. (2012) concluded that if religion is made salient in men's minds, their displays of endurance (in addition to their willingness to delay gratification) during a subsequent task are reduced.

### 1.1. Overview of the present study and predictions

Here, we sought to replicate the findings reported by McCullough et al. (2012), following the methods and analyses they used in their *Experiment 3* as closely as possible, with an added sex of experimenter factor to shed further light on the effects of religious primes on men's hand grip endurance. We also thought it was worthwhile to examine the robustness of McCullough et al.'s (2012) *Experiment 3* findings in light of recent failures to replicate other experimental results that have used similar implicit primes (e.g., Carlin & Standing, 2013; Doyen, Klein, Pichon, & Cleeremans, 2012; Pashler, Coburn, & Harris, 2012). We expected a sex by condition interaction, revealing that men who unscrambled sentences embedded with religious content would demonstrate shorter maximum endurance times than would their peers in the control condition. In their original experiment, McCullough et al. (2012) used only male experimenters in an effort to obtain experimental control. However, because experimenter sex has been found to influence peoples' performance in other arenas (e.g., Ronay & von Hippel, 2010), in the experiment reported here, we used both male and female experimenters in hopes of better characterizing the motivational basis of the behavioral changes that the religious prime produced in men. We reasoned that if the effects of the religious prime in down-regulating hand-grip endurance manifested themselves only among male participants who had been paired with male experimenters, we would be in a position to conclude that sexual primes reduced displays of physical endurance for the purpose of signaling physical prowess to other males, but not for the purpose of signaling physical prowess to women.

## 2. Method

### 2.1. Participants

Participants were 168 women and 159 men with a mean age of 19.03 ( $SD = 2.05$ ) recruited from introductory psychology courses at the University of Miami (other findings resulting from this data collection effort, not relevant to the present paper, appear in Hone & McCullough, 2012). Participants received \$7.00 in addition to a small amount of course credit for their participation. We collected data on an additional 37 participants but we excluded them from analyses (consistent with the exclusion criteria from McCullough et al., 2012, *Experiment 3*) for the following reasons: Either they completed the hand grip task incorrectly (which included opting out due to hand injuries), or they completed the priming task incorrectly (which included not following directions). Consistent with the methods in McCullough et al. (2012)'s *Experiment 3*, we randomly assigned participants to either the religious prime condition or the control condition, blocking on sex. Table 1 displays the cell sizes for male and female participants in the religious prime and control conditions who were run by male and female experimenters.

**Table 1**

Cell sizes (data collected by male and female experimenters).

Participant Sex	Experimenter Sex	Condition	Frequency	Percent
Female	Female	Control	32	43.2
		Religious	42	56.8
	Male	Control	43	45.7
		Religious	51	54.3
Male	Female	Control	42	47.7
		Religious	46	52.3
	Male	Control	39	54.9
		Religious	32	45.1

### 2.2. Procedure

As closely as possible, we followed the experimental details outlined in McCullough et al. (2012)'s *Experiment 3*, although we used different undergraduate experimenters (four male; seven female) to run sessions. During the experiment, we first recorded participants' grip strength (maximum voluntary contraction; Little & Johnson, 1986) using a hand-grip dynamometer (LaFayette Instruments, Lafayette, IN, USA; model 78010). We then randomly assigned participants (blocking on sex) to one of two scrambled sentence tasks (Shariff & Norenzayan, 2007; E.L. Uhlmann, personal communication, September 25, 2008). In the religious condition, participants were asked to unscramble 20 sentences, 10 of which contained a word related to religion (e.g., "divine" or "sacred"). In the control condition, participants were asked to unscramble 20 sentences that did not consistently prime any concept. We then measured how long participants could maintain a force equal to 70% of their maximum voluntary contraction, also known as maximum endurance time (Little & Johnson, 1986). Following McCullough et al. (2012)'s *Experiment 3*, we used natural log-transformed maximum endurance time as our dependent variable. Also in line with McCullough et al. (2012), experimenters remained ignorant of participants' condition assignments throughout the experiment. All scrambled sentence tasks were distributed to participants in sealed packets, participants were instructed to place packets in a box upon completion, and experimenters did not handle data until participants had left the laboratory. The data analyzed for this experiment, and for *Experiment 3* in McCullough et al. (2012) can be accessed at <https://osf.io/y6mi9/>.

### 2.3. Analyses

First, we ran a two (sex of participant: male or female) by two (sex of experimenter: male or female) by two (condition: religious prime or control) full-factorial ANOVA on the natural log-transformed maximum endurance times. Upon finding no significant main effect or interactions involving experimenter sex, we reduced the model to a two (sex of participant: male or female) by two (condition: religious prime or control) ANOVA with two main effects and an interaction of sex and condition. We explored these results with another two-way ANOVA in which we excluded all participants whose experimental sessions were conducted by female experimenters to make our results more comparable to those in McCullough et al. (2012). Finally, we conducted a multilevel model to examine possible experimenter effects (Raudenbush & Bryk, 2002 see Supplementary Online Material, available on the journal's website at [www.ehbonline.org](http://www.ehbonline.org)).

### 2.4. Power analysis

McCullough et al. (2012)'s *Experiment 3* comprised 167 participants, including 78 men (see Table 2). The effect size of the difference between mean natural log-transformed maximum endurance time of men in the religious prime condition versus the control condition was Cohen's  $d = .50$ . Our replication comprised  $N = 327$  participants, including  $n = 159$  men (see Table 1). Assuming an effect size of  $d = .50$ ,  $n = 159$  men, and  $p < .05$  (two-tailed), our power to detect a simple effect of religious

**Table 2**

Cell sizes: McCullough et al. (2012).

Participant Sex	Condition	Frequency	Percent
Female	Religious	40	48.8
	Control	42	51.2
Male	Religious	39	50
	Control	39	50

prime on natural log-transformed maximum endurance time for men was .88. For only the subsample of participants ( $N = 165$ , including  $n = 71$  men) that were run by male experimenters, the comparable power was .54.

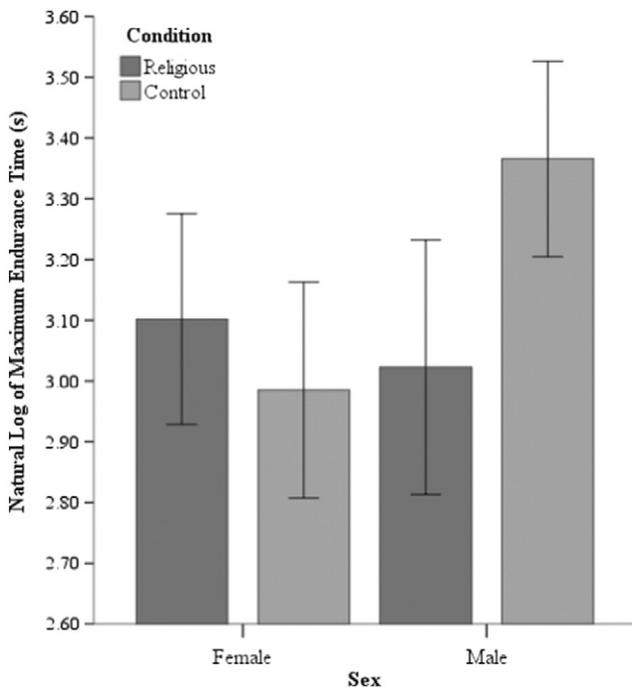
### 3. Results

#### 3.1. Three-way ANOVA

The dependent variable, maximum endurance time, was not normally distributed, so following McCullough et al. (2012), we natural log-transformed the values. We then ran a two (sex of participant: male or female) by two (sex of experimenter: male or female) by two (condition: religious prime or control) full-factorial ANOVA on the natural log-transformed maximum endurance times. We did not find (1) a main effect of experimenter sex,  $F(1,318) = 1.04$ ,  $p = .31$ ; (2) an interaction between experimenter sex and participant sex,  $F(1,318) = .01$ ,  $p = .92$ ; (3) an interaction between experimenter sex and condition,  $F(1,318) = .04$ ,  $p = .85$ ; or (4) a three-way interaction between participant sex, experimenter sex, and condition,  $F(1,318) = 1.02$ ,  $p = .31$ . Consequently, we removed the effects involving experimenter sex from the model and ran a two-way ANOVA.

#### 3.2. Two-way ANOVA

We ran a two (sex of participant: male or female) by two (condition: religious prime or control) full-factorial ANOVA on the natural log-transformed maximum endurance times, ignoring the sex of the experimenter. Fig. 1 depicts the means and 95% confidence intervals for men

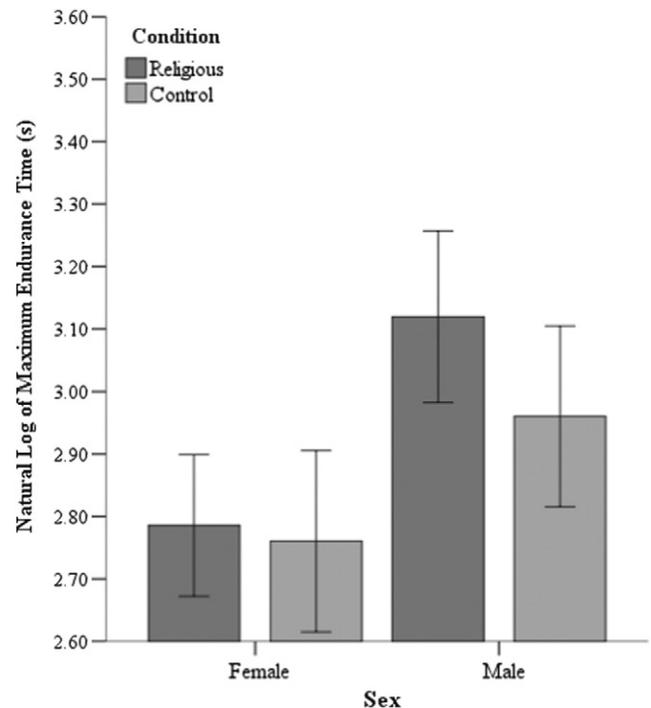


**Fig. 1.** Group mean differences: McCullough et al. (2012)'s Experiment 3. Note. Error bars = 95% confidence intervals.

and women in the two experimental groups from McCullough et al. (2012) and Fig. 2 shows the means and 95% confidence intervals for men and women in the two experimental groups from the present replication. Consistent with McCullough et al. (2012), we found no main effect for condition,  $F(1,322) = 1.90$ ,  $p = .17$ : Participants in the religious prime condition ( $M = 2.94$ ,  $SD = .60$ ) did not perform differently on the endurance task than their did peers in the control condition ( $M = 2.86$ ,  $SD = .65$ , effect size  $d = .12$ , 95% CI:  $-.21$ ;  $.06$ ). Although McCullough et al. (2012) found a small and statistically non-significant tendency for men to evince greater endurance times than did women,  $p = .09$ ,  $d = +.26$  (which McCullough et al. erroneously labeled as  $d = -0.26$ ), the sex difference in the present experiment was slightly larger and statistically significant,  $F(1,322) = 15.40$ ,  $p < .001$ ,  $d = +.43$ . Men ( $M = 3.04$ ,  $SD = .63$ ) had higher natural log-transformed maximum endurance times than did women ( $M = 2.77$ ,  $SD = .59$ , effect size  $d = .43$ , 95% CI:  $-.40$ ;  $-.13$ ). We did not replicate McCullough et al.'s (2012) significant interaction of sex and condition,  $F(1,322) = .98$ ,  $p = .32$ : That is, we did not find that men in the religious condition ( $M = 3.12$ ,  $SD = .61$ ) performed worse on the grip task from their peers in the control condition ( $M = 2.96$ ,  $SD = .65$ , effect size  $d = .25$ , 95% CI:  $-.36$ ;  $.04$ ).

#### 3.3. Two-way ANOVA excluding data collected by female experimenters

To further address potential effects of experimenter sex, and in an attempt to more directly replicate the results of McCullough et al. (2012), we re-ran the two-way ANOVA excluding participants whose sessions were conducted by female experimenters, which left us with a subsample of 165 participants, 71 of whom were men (see Table 3). Here, too, we were again unable to replicate Experiment 3 of McCullough et al. (2012). That is, we did not find an effect of condition,  $F(1,160) = .52$ ,  $p = .47$ , effect size  $d = .08$ ; though we did find an effect of sex  $F(1,160) = 2.46$ ,  $p = .009$ . Men ( $M = 2.99$ ,  $SD = .63$ ) had higher natural log-transformed maximum endurance times than did women ( $M = 2.75$ ,  $SD = .56$ , effect size  $d = .40$ , 95% CI:  $-.43$ ;  $-.06$ ). We also did not replicate the sex by condition effect,  $F(1,160) = .003$ ,  $p = .96$ . Fig. 3 shows the means and 95% confidence intervals for men



**Fig. 2.** Group mean differences (data collected by male and female experimenters). Note. Error bars = 95% confidence intervals.

**Table 3**  
Cell sizes (data collected by male experimenters only).

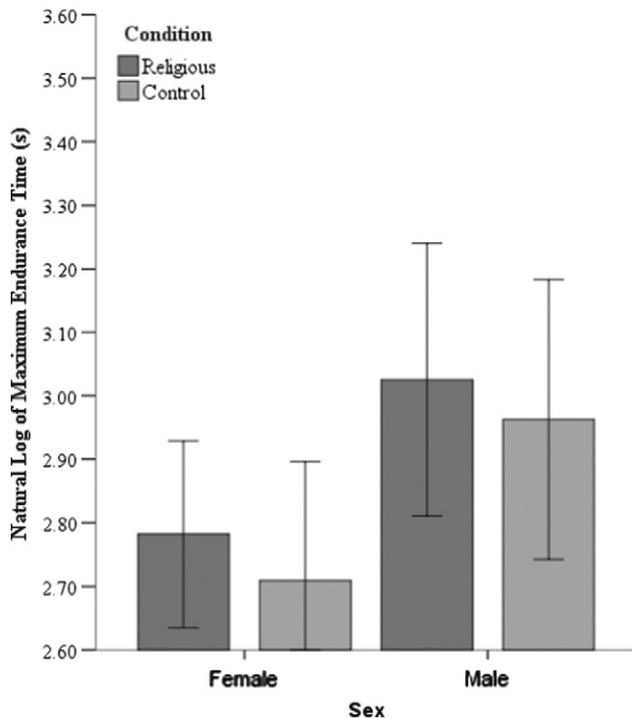
Participant Sex	Condition	Frequency	Percent
Female	Control	43	45.7
	Religious	51	54.3
Male	Control	39	54.9
	Religious	32	45.1

and women in the two experimental groups who were run by male experimenters.

#### 4. Discussion

Despite obtaining adequate power to detect an effect of religious cognition (manipulated via a scrambled sentence task) on a (presumably) sexually selected male characteristic (maximum endurance time), we failed to replicate the findings of McCullough et al. (2012)'s Experiment 3. This adequately powered failure to replicate suggests that the effect of implicit religious priming on men's (but not women's) hand grip endurance may not be real, and therefore, that the results in McCullough et al. (2012) were plausibly due to Type I error. These results join several other attempts at exactly replicating previously published evidence that behavior can be changed with social priming (e.g., Carlin & Standing, 2013; Doyen, Klein, Pichon, & Cleeremans, 2012; Pashler, Coburn, & Harris, 2012).

Aside from the possibility that it is simply not true that implicit religious priming influences men's hand grip endurance (and that McCullough et al.'s, 2012 results on this topic were a false positive), it is also possible that the effect of religious priming on hand grip endurance was dependent on unique aspects of McCullough et al.'s (2012) Experiment 3 that we have been unable to empirically identify here. In the present replication, we tried as closely as possible to directly replicate McCullough et al.'s (2012) methods, but there were two factors we could not control: (1) the subjects themselves; and (2) the experimenters who ran their sessions.



**Fig. 3.** Group mean differences (data collected by male experimenters only). Note. Error bars = 95% confidence intervals.

It is conceivable that there were systematic differences between the subjects we ran and those that McCullough et al. (2012) ran, but in both the original experiment and the present experiment, sample sizes were reasonably large and subjects were drawn from the same subject pool, so this explanation seems the less plausible of the two. What seems more plausible is that the two experimenters (two males who, it seems to us, had an above-average level of athleticism and social poise) who ran participants in McCullough et al.'s (2012) Experiment 3 elicited different levels of performance from their subjects than did the 11 experimenters (seven female and four males) who ran the present experiment. Experimenter effects might account not only for our failure to replicate McCullough et al.'s (2012) experimental findings, but also for the generally reduced maximum endurance times in our experiment relative to the original study. This suspicion is supported indirectly by the fact that we did find significant between-experimenter variance in the effects of participant sex on hand grip endurance, even though we could not attribute this variance to experimenters' sex (cf. Ronay & von Hippel, 2010). Note, however, that we did not find between-experimenter variance in the condition by sex interaction, which was the hypothesis of focal interest in both McCullough et al. (2012) and here (see Supplementary Online Material, available on the journal's website at [www.ehbonline.org](http://www.ehbonline.org)).

Regrettably, we have no objective measures of our experimenters' (or McCullough et al.'s, 2012 experimenters') social dominance, athleticism, physical formidability, or attractiveness, so we are unable to test these intuitions systematically. Nevertheless, it is generally well-accepted both as a methodological principle and as a matter of theoretical interest that characteristics of experimenters can influence subjects' performance in systematic ways (Harari et al., 1970), although we wonder how often researchers take this possibility seriously when they run laboratory experiments in which subjects interact with experimenters whose effects are presumed (perhaps without adequate license for doing so) to be negligible.

#### 4.1. Conclusion

Here, we were unsuccessful in our efforts to directly replicate the effect of religious priming on maximum endurance time on a hand-grip task that McCullough et al. (2012) reported. Social priming experiments – indeed, experiments of all types – can be difficult to replicate due to Type I errors or experimenter error (Carlin & Standing, 2013; Pashler et al., 2012). It is possible that the findings reported by McCullough et al. (2012) were a result of a Type I error. It is also possible that our failure to replicate McCullough et al. (2012) was due to uncontrolled variations in experimenter effects, particularly as participants in the McCullough et al. (2012) study interacted with different experimenters than the participants run through the replication we report here (see Supplementary Online Material, available on the journal's website at [www.ehbonline.org](http://www.ehbonline.org)). Until further evidence becomes available, we think the claim that religious priming reduces men's hand grip endurance (McCullough et al., 2012) should be regarded with circumspection.

#### Supplementary Materials

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.evolhumbehav.2014.08.007>.

#### Acknowledgments

We thank the Evolution and Human Behavior Laboratory research assistants for their data collection efforts.

## References

- Archer, J. (2009). Does sexual selection explain human sex differences in aggression? *Behavioral and Brain Sciences*, 32, 249–311, <http://dx.doi.org/10.1017/S0140525X09990951>.
- Carlin, S. P., & Standing, L. G. (2013). Is intelligence enhanced by letter priming? A failure to replicate the results of Ciani and Sheldon (2010). *Psychological Reports*, 112(2), 533–544, <http://dx.doi.org/10.2466/04.03.PR0.112.2.533-544>.
- Daly, M., & Wilson, M. (2005). Carpe diem: Adaptation and devaluing the future. *The Quarterly Review of Biology*, 80(1), 55–60 (Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/15884736>).
- Doyen, S., Klein, O., Pichon, C. -L., & Cleeremans, A. (2012). Behavioral priming: It's all in the mind, but whose mind? *PLoS one*, 7(1), e29081, <http://dx.doi.org/10.1371/journal.pone.0029081>.
- Harari, H., Jacobson, E. A., Peter, M., Nelson, T. A., Sattler, V. R., Sattler, M., et al. (1970). Racial "experimenter effects" in experimentation, testing, interviewing, and psychotherapy. *Psychological Bulletin*, 73(2), 137–160.
- Hawkes, K. (1991). Showing off. *Ethology and sociobiology*, 12, 29–54, [http://dx.doi.org/10.1016/0162-3095\(91\)90011-E](http://dx.doi.org/10.1016/0162-3095(91)90011-E).
- Hone, L. S. E., & McCullough, M. E. (2012). 2D:4D ratios predict hand grip strength (but not hand grip endurance) in men (but not in women). *Evolution and Human Behavior*, 33(6), 780–789, <http://dx.doi.org/10.1016/j.evolhumbehav.2012.07.003>.
- Kirby, K. N., & Maraković, N. N. (1996). Delay-discounting probabilistic rewards: Rates decrease as amounts increase. *Psychonomic Bulletin & Review*, 3(1), 100–104, <http://dx.doi.org/10.3758/BF03210748>.
- Little, M. A., & Johnson, B. R. (1986). Grip strength, muscle fatigue, and body composition in nomadic Turkana pastoralists. *American Journal of Physical Anthropology*, 69, 335–344, <http://dx.doi.org/10.1002/ajpa.1330690306>.
- McCullough, M. E., Carter, E. C., DeWall, C. N., & Corrales, C. M. (2012). Religious cognition down-regulates sexually selected, characteristically male behaviors in men, but not in women. *Evolution and Human Behavior*, 33(5), 562–568, <http://dx.doi.org/10.1016/j.evolhumbehav.2012.02.004>.
- Pashler, H., Coburn, N., & Harris, C. R. (2012). Priming of social distance? Failure to replicate effects on social and food judgments. *PLoS one*, 7(8), e42510, <http://dx.doi.org/10.1371/journal.pone.0042510>.
- Pawlowski, B., Atwal, R., & Dunbar, R. I. M. (2008). Sex differences in everyday risk-taking behavior in humans. *Evolutionary Psychology*, 6, 29–42.
- Putnam, R. D., & Campbell, D. E. (2012). *American grace: How religion divides and unites us*. Simon and Schuster.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods, Vol. 1*, Sage.
- Ronay, R., & von Hippel, W. (2010). The presence of an attractive woman elevates testosterone and physical risk taking in young men. *Social Psychological and Personality Science*, 1(1), 57–64, <http://dx.doi.org/10.1177/1948550609352807>.
- Shariff, A. F., & Norenzayan, A. (2007). God is watching you. *Psychological Science*, 18(9), 803–809.
- Shih, Y. -C. (2007). Glove and gender effects on muscular fatigue evaluated by endurance and maximal voluntary contraction measures. *Human Factors*, 49, 110–119, <http://dx.doi.org/10.1518/001872007779598091>.
- Silverman, I. W. (2003). Gender differences in delay of gratification: A meta-analysis. *Sex Roles*, 49, 451–463.
- Weeden, J., Cohen, A. B., & Kenrick, D. T. (2008). Religious attendance as reproductive support. *Evolution and Human Behavior*, 29(5), 327–334, <http://dx.doi.org/10.1016/j.evolhumbehav.2008.03.004>.
- Weeden, J., & Kurzban, R. (2013). What predicts religiosity? A multinational analysis of reproductive and cooperative morals. *Evolution and Human Behavior*, 34, 440–445, <http://dx.doi.org/10.1016/j.evolhumbehav.2013.08.006>.
- Wilson, M., & Daly, M. (2004). Do pretty women inspire men to discount the future? *Proceedings of the Royal Society B*, 271, S177–S179, <http://dx.doi.org/10.1098/rsbl.2003.0134>.