

Replication of the Superstition and Performance Study by Damisch, Stoberock, and Mussweiler (2010)

Robert J. Calin-Jageman and Tracy L. Caldwell

Dominican University, River Forest, IL, USA

Abstract. A recent series of experiments suggests that fostering superstitions can substantially improve performance on a variety of motor and cognitive tasks (Damisch, Stoberock, & Mussweiler, 2010). We conducted two high-powered and precise replications of one of these experiments, examining if telling participants they had a lucky golf ball could improve their performance on a 10-shot golf task relative to controls. We found that the effect of superstition on performance is elusive: Participants told they had a lucky ball performed almost identically to controls. Our failure to replicate the target study was not due to lack of impact, lack of statistical power, differences in task difficulty, nor differences in participant belief in luck. A meta-analysis indicates significant heterogeneity in the effect of superstition on performance. This could be due to an unknown moderator, but no effect was observed among the studies with the strongest research designs (e.g., high power, *a priori* sampling plan).

Keywords: superstition, luck, self-efficacy, replication

Can superstitions actually improve performance? Damisch, Stoberock, and Mussweiler (2010) reported a striking experiment in which manipulating superstitious feelings markedly increased golfing ability. Participants attempted 10 putts, each from a distance of 100 cm. Some participants were primed for superstition prior to the task by being told “Here is the ball. So far it has turned out to be a lucky ball.” Controls were simply told “This is the ball everyone has used so far.” Remarkably, this manipulation produced a substantial increase in golf performance: Controls made 48% of putts while superstition-primed participants made 65% of putts ($d = 0.83$, 95% CI [0.05, 1.60]).

This simple experiment suggests a major revision to our notions of superstition. The very definition of a superstition is a belief that is “irrational” which arises due to a “false conception of causation” (Merriam-Webster, 2013). Indeed, there has been a long scientific tradition of pointing out the fundamental lack of efficacy of superstitious behavior (e.g., Galton, 1872). The prevalence of superstitious behavior has thus been classically explained as an effect of confirmation bias rather than a true association with reinforcing outcomes (Skinner, 1948). In contrast, the results from Damisch et al. (2010) suggest that superstitions about one’s own behavior can be efficacious. If true, this class of superstition is not completely irrational and the prevalence of such behaviors could be explained by their strong positive consequences. Both psychologists and the general public have been quick to recognize the importance of this finding. The original report has been cited 55 times (Google Scholar, 2013), was covered extensively in the popular press at the time of publication (e.g., Doheny, 2010;

Hutson, 2010), and has even become part of the sales pitch for an online website selling lucky charms (<http://www.good-luck-gifts.com>, n.d.).

In support of their findings, Damisch et al. (2010) reported three successful conceptual replications. In addition, a dissertation by Damisch (2008) reports an additional two successful conceptual replications. This work is summarized in Table 1 and Figure 1. Integration across results indicates an overall effect size that is at least moderate and possibly very large (unbiased $d = 0.82$, 95% CI [0.53, 1.11], white diamond in Figure 1).

While these results suggest a robust and powerful effect of superstition on performance, conceptual replications by others show mixed results. Lee, Linkenauger, Bakdash, Joy-Gaba, and Proffitt (2011) found that golfers told they were using a famous golfer’s putter performed substantially better on a putting task than controls. In contrast, Aruguete, Goodboy, Jenkins, Mansson, and McCutcheon (2012) found that superstitions related to prayer are not effective at improving performance on a reasoning test. In an additional experiment, priming participants to think about their religious beliefs also failed to improve scores on a reasoning test relative to controls.

The contrasting results of these conceptual replications could be due to a number of factors. It could be that only some types of superstitions are efficacious, perhaps those related to luck rather than religion. Another possibility is that superstition can affect performance on only some types of tasks. Given the uncertainty, it seemed important as a first step to directly confirm the replicability of the original finding. Here we report two high-powered, precise

Table 1. Summary of studies examining the effect of superstition on performance

Study	Superstition Prime	Dependent Variable	<i>n</i> /group	Power if <i>d</i> = 0.83
Damisch et al. (2010)				
Experiment 1	Experimenter labeled ball lucky	Golf task	14	0.55
Experiment 2*	Experimenter wished good luck	Motor-dexterity task	17	0.64
Experiment 3	Presence of self-selected lucky charm	Memory game	20	0.71
Experiment 4	Presence of self-selected lucky charm	Anagram task	15	0.58
Damisch (2008)				
Study 2	Experimenter made shot and labeled ball lucky	Golf task	14	0.55
Study 5	Subliminal priming for word “luck”	Tetris	14	0.55
Lee et al. (2011)				
Experiment 1	Told using a professional golfer’s putter	Golf task	20	0.71
Aruguete et al. (2012)				
Experiment 1	Religious beliefs scale	Verbal reasoning test	71	0.99
Experiment 2*	Wrote and prayed for success	Verbal Reasoning test	53	0.99
This manuscript				
Experiment 1	Experimenter labeled ball lucky	Golf task	58	0.99
Experiment 2	Drew ball with clover and experimenter labeled lucky	Golf task	54	0.99

Notes. For *n*/group, smaller group size reported if group sizes were uneven. Power calculated for an effect size of 0.83, the unbiased overall effect size calculated over the six studies by Damisch (2008) and Damisch et al. (2010). *Additional groups run beyond control and superstition prime, but only 1 contrast selected for meta-analysis.

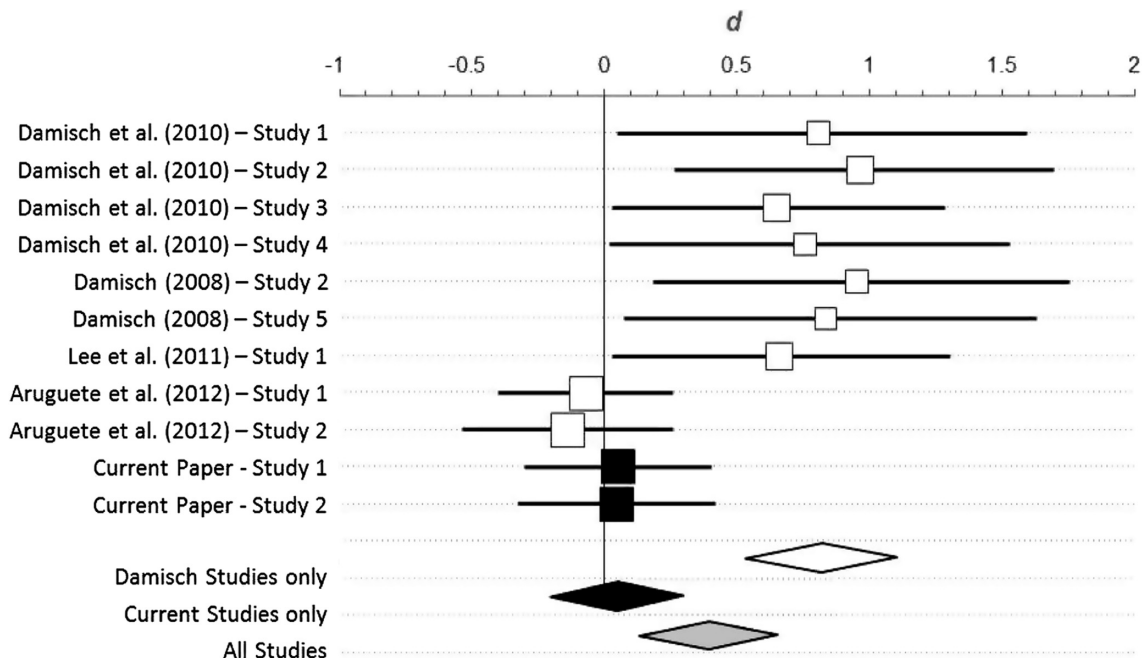


Figure 1. Meta analysis of the effects of superstition on performance. The location of each square represents the observed effect size of a single study. The 95% CI of the effect size is represented by the line extending from the square, and relative sample size is represented by the area of the square. Studies conducted prior to this one are shown with white squares; the two studies reported in this manuscript are shown in black. The diamonds represent unbiased effect sizes estimates over groups of studies, with the center of the diamond marking the point estimate for effect size and the width of the diamond covering the 95% CI. The first overall estimate (white diamond) is for the six studies conducted by Damisch (2008) and Damisch et al. (2010). The black diamond represents the overall effect size estimate from the two studies reported in this manuscript. The gray diamond is for all studies, but note that significant heterogeneity of effect sizes was evident ($Q(10) = 26.5, p = .003$). This figure was created using ESCI (Cumming, 2011).

replications of the golf and superstition experiment (Study 1) from Damisch et al. (2010). We focused on this study because it is simple (just two groups), involves no idiomatic language, applies to the general population, involves equipment which can be precisely matched to the original experiment, and has a large effect size.

Experiment 1: Direct Replication

We matched both the materials and procedures of the target study as precisely as possible. This was facilitated by the gracious cooperation of Lysann Damisch (personal communications, 2012–2013) who provided detailed feedback based on a review of our materials and a video of a training session.

The replication was registered in advance of data collection on the Open Science Framework. All materials, data, video of the procedure, and the preregistered design are available at <http://osf.io/fsadm/>. We report all data exclusions, manipulations, and measures, and how we determined our sample sizes.

Method

Participants

The target study's data came from a convenience sample of German undergraduates, with 57% females and 43% males. Left-handers and psychology majors were excluded (Damisch, personal communication). No compensation was provided. Eighty percent of the participant pool believed in superstition. This was based on responses to a single item "How much do you believe in good luck and bad luck" rated on a scale from 1 (= *not at all*) to 9 (= *very much*), with responses greater than 3 counted as belief in superstition (Damisch, personal communication).

We collected a convenience sample from biology laboratory classes at a private comprehensive university in the United States. The biology classes we targeted were open to non-majors and most fulfilled general education requirements, leading to enrollment from a wide range of majors. Moreover, science majors within our university exhibit similar levels of superstition compared to the participants in the target study (79%, or 27 out of 34 responses to the same item delivered as an online questionnaire to Chemistry, Natural Science, and Mathematics majors at the same university).

We did not exclude left-handers because we did not know about this criterion until after our sampling plan was developed. However, we fortuitously targeted classes with relatively few psychology majors, and tracked major so that these participants could be excluded post hoc. Participants were compensated with an entry into a drawing to receive a \$25 Amazon gift card, with odds of winning set at 1:20.

We planned to sample at least 42 but no more than 91 participants per group. The minimum was set to provide 0.95 power for the overall average effect size (0.82) across the six prior superstition and performance studies (Damisch et al., 2008, 2010), the maximum to provide similar power for the lower bound of the 95% CI for the effect (0.53). We collected data until our minimum target had been exceeded and our participant pool was depleted, yielding data from 58 controls and 66 superstition-activated participants (power at 0.83 even for the lower-bound effect size estimate). Our sample consisted of 90 females (73%), 28 males (23%) and 6 participants who did not report their gender (5%). No participants were excluded from initial analysis. Although this overrepresents females relative to the target study, the effects of superstition on performance have been demonstrated with all-female samples (Damisch et al., 2010, Study 2).

Materials

Three female research assistants collected all the data for this study. To ensure smooth and even-handed administration of the experiment, each assistant memorized an experimental script and completed at least five practice sessions prior to collecting data. None of the research assistants had read the target article, but were informed that the manipulation could enhance, impair, or have no effect on performance.

Two approximately identical research rooms were used for data collection. Each contained a personal computer with the monitor surrounded by a study carrel for privacy. Each also had an open space for the putting task. The floor was covered with office-grade brown wall-to-wall carpeting.

Participants completed a computerized questionnaire at the beginning to record consent, gender, major, and school year. In addition, text instructions explained that they would complete a golf task because adapting to a new task is a good predictor of future success. This cover story was provided by Damisch (personal communication).

We acquired the same executive putting set (Shivam Putter Set in Black Travel Zipper Pouch, see source list) used in the original research. The set consists of a metal putter, two standard white golf balls, and a square wooden target with an omega-shaped cutout. We replaced the putter, however, with a similar model made for both left- and right-handed putters (Quolf Two-Way Putter), to accommodate left-handed participants.

Damisch et al. (2010) used a putting distance of 100 cm. In a pilot test, we found that students in our undergraduate population are too good at this task, due either to more golf experience or to a slower "green speed" for the carpeting in our research rooms. Controls ($n = 8$) averaged 8.25/10, considerably higher than the 4.8/10 reported for controls in the original study. We therefore moved the target back to 150 cm to equate difficulty. In a second round of pilot testing at this distance, controls ($n = 19$) averaged 5.9/10, much closer to the original study. We used this longer putting distance to achieve similar task difficulty.

The target itself was placed 100 cm from the wall (Damisch, personal communication), and the starting point for the ball was marked with tape.

To ensure and measure the quality of the replication, we added a quality-control task and a manipulation check via a computerized post golfing task questionnaire. Participants were asked “What did the researcher say to you as she handed you the golf ball?” Participants in the lucky condition passed if they mentioned the word *luck* (or any of its variants); participants in the control condition passed if they failed to mention the word *luck* (or any of its variants).

Then, participants completed a two-item manipulation check: “Before starting this task, I believed that the golf ball assigned to me was lucky” and “Now that I have completed this task, I believe that the golf ball assigned to me is lucky.” Responses were made on a Likert scale from 1 (= *strongly disagree*) to 5 (= *strongly agree*). Note that these manipulation checks were retrospective. Responses could thus be contaminated by their experience on the golf task. This order was used, however, to avoid altering the original protocol. Furthermore, pilot testing suggested that these measures would still elicit the expected group difference in feelings of luck, $t(41) = 2.23$, $p = .031$, $d = 0.66$.

Score sheets were created in advance for each participant with random assignment to the control or superstition-primed group via a random number generator. Score sheets were then printed, and placed in the research rooms for the research assistants to use in sequential order. Condition was indicated on the score sheet as a “C” or “L” to avoid priming participants should they glance at the score sheet.

Procedure

Participants were recruited during down-time in their laboratory class sessions. Volunteers were escorted to the research room one at a time by a research assistant. Upon arrival, the researcher asked the participant to complete the initial portion of the computerized questionnaire, including informed consent, demographics, cover story, and task explanation. The researcher then explained the task again and handed the participant the golf ball, saying either “Here is the ball. So far it has turned out to be a lucky ball” (superstition-activated group) or “This is the ball everyone has used so far” (control group).

Participants then completed the golf task (10 putts). After each shot, the researcher stated “Ok, now for shot X” where X was the next shot. No other feedback was given. After the golf task, the participants completed the quality-control task and manipulation check. The research assistant stood on the other side of the room during this task.

Differences From the Original Study

We conducted a faithful replication of the golf and superstition study by Damisch et al. (2010). The only differences are that we:

- recruited US college students rather than German college students.

- administered the experimental script in English rather than German (but using the translation provided by Damisch et al. (2010) for the key manipulation).
- recruited a somewhat higher proportion of women.
- collected data with three female research assistants rather than one.
- used a putting distance of 150 cm rather than 100 cm to equalize task difficulty for our population.
- included left-handed golfers as well as right-handed golfers.
- added a quality-control task and manipulation check to the end of the protocol.

Most of these differences are not substantive, with the possible exception of cultural differences between undergraduates from Germany and the US. However, similar results have been reported with students living in the US. (Lee et al., 2011), and our participants were well-matched in terms of their belief in good and bad luck.

Analysis

As in the original report, differences in performance between the superstition and control groups were analyzed with an independent samples *t* test. Effect sizes are estimated using Cohen’s *d*. We also report confidence intervals for group differences. Estimates of power were calculated with PS Power and Sample Size Software for Windows (Dupont & Plummer, 1998)

Results and Discussion

We did not observe a strong impact of superstition on golf performance (Table 2). The superstition-activated group performed just 2% better than the control group (compared to 35% improvement in the target study). This difference did not reach statistical significance, $t(122) = 0.29$, $p = .77$.

Participants in the superstition-activated group retrospectively reported themselves to have felt luckier at the start of the golf task compared to those in the control group, $t(115.4) = 4.28$, $p = .00004$. This feeling of luck was also evident after the golf task was complete, $t(112) = 2.02$, $p = .045$. Despite the successful manipulation checks, we did find that many participants in the superstition-activated group failed the quality-control task we designed. Specifically, when asked, “What did the experimenter say when she handed you the golf ball?” only 42 of 66 (63%) participants mentioned “luck.” Excluding the participants who failed this task still preserved strong power for the analysis (0.98), but the group difference remained very small (3.6% improvement) and did not reach statistical significance, $t(98) = 0.40$, $p = .69$.

Debriefing provided some clues as to why so many participants in the superstition-activated group failed the quality-control task. Some participants reported that they believed the mention of luck by the experimenter was “off script” and had not wanted to mention it for fear of getting the experimenter in trouble. Thus, some participants

Table 2. Effects of superstition on golf performance

Measure	Control <i>M</i> (<i>SD</i>)	Superstition <i>M</i> (<i>SD</i>)	Mean difference <i>M</i> [95% CI]	Effect size <i>d</i>
Golf putts made	4.62 (2.13)	4.73 (1.96)	0.11 [– 0.62, 0.83]	0.05
Manipulation checks:				
Felt lucky prior to task	1.64 (0.93)	2.55 (1.36)	0.91 [0.50, 1.32]	0.47
Felt lucky after task	2.02 (1.26)	2.48 (1.30)	0.47 [0.10, 0.93]	0.36
Quality control:				
Recalled prompt	58/58 (100%)	42/66 (63%)		
Golf putts made	4.62 (2.13)	4.79 (1.97)	0.17 [– 0.66, 0.99]	0.08

Notes. Control group $n = 58$, superstition-activated group $n = 66$. Note that manipulation checks were retrospective ratings made *after* the task. Under quality control, putts made are only for those participants who passed the quality control task.

may have failed this task not due to poor impact but due to highly credulous responses to the manipulation. Exploratory analysis provided some evidence consistent with this interpretation; those in the superstition-activated group who failed the quality-control task actually reported slightly *higher* feelings of luck than those who passed (e.g., $M = 2.88$, $SD = 1.29$ for the 14 participants who failed the task; $M = 2.36$, $SD = 1.38$ for the 42 participants who passed the task, though this difference is not statistically significant, $t(64) = 1.50$, $p = .14$).

We conducted exploratory analyses to try to uncover an effect of superstition on performance. We excluded psychology majors, checked for an interaction by gender, and checked for an interaction by research assistant. No significant effects were observed (see Supplementary Table S1).

Experiment 2: Higher Impact Replication

Although our replication attempt succeeded in having high power and demonstrable impact, we wondered if a stronger superstition prime would produce the expected effect.

Method

Methods were the same as above but with the following modifications.

Participants

We altered our sampling plan to target the general university population to ensure that the results were not idiosyncratic to students enrolled in biology courses. The sample was recruited by advertising on campus bulletin boards and on campus. For this study, participants signed up for appointments and arrived at the research room on their own. Participants were compensated with an experimental participation voucher that could be redeemed in some classes toward course credits.

We collected data for 113 participants, halting data collection when our minimum target was exceeded and only one week remained before the deadline for manuscript submission. One participant in the lucky condition requested at the end of the experiment that his or her data be withdrawn from analysis. Another failed the quality-control task and was removed. Thus, our final sample consisted of 111 participants (28 males and 83 females); these were randomly assigned to either the control ($n = 54$) or superstition-activated ($n = 66$) conditions.

Materials

To enhance impact, participants selected their ball from a velour sack containing eight golf balls: Four regular and four emblazoned with a green clover (Shamrock Golf Ball, see source list). The experimenter's prompt was also enhanced: "This is the ball you will use" for the control group versus "Wow! You get to use the lucky ball" for those in the superstition-activated group.

Table 3. Effects of enhanced superstition activation on golf performance

Measure	Control <i>M</i> (<i>SD</i>)	Superstition <i>M</i> (<i>SD</i>)	Mean difference <i>M</i> [95% CI]	Effect size <i>d</i>
Golf putts made	4.02 (2.20)	4.12 (2.01)	0.10 [– 0.69, 0.90]	0.05
Manipulation checks:				
Felt lucky prior to task	2.02 (0.93)	2.86 (1.4)	0.85 [0.38, 1.30]	0.68
Felt lucky after task	2.07 (0.95)	2.46 (1.28)	3.47 [– 0.04, 0.81]	0.34
Quality control:				
Recognized ball	54/54 (100%)	66 of 67 (99%)		

Notes. Control group $n = 54$, superstition-activated group $n = 66$ (one participant who failed quality-control task excluded from data analysis). Note that manipulation checks were retrospective ratings made *after* the task.

To explore possible moderators, we added a measure of belief in luck. This was the same measure described earlier that Damisch et al. had used to measure belief in luck in their participant population (Damisch, personal communication). This item, like the others, was administered via a computerized questionnaire after the golf task was completed.

The quality-control task was modified to a recognition task: Participants were shown an image of both the regular ball and the “lucky” ball and asked to choose which they had received. To avoid contaminating other responses by showing both conditions the “lucky” ball, this task was moved to the end of the questionnaire.

Results and Discussion

We did not observe an impact of superstition on performance (see Table 3). Participants in the superstition-activated group scored just 2.5% higher than those in the control group, a nonsignificant difference, $t(109) = 0.26$, $p = .80$.

Our failure to replicate was not due to insufficient impact, as this study produced an even larger difference in participants’ retrospectively reported feelings of luck before the golf task, $t(96) = 3.65$, $p = .004$. The difference in ratings of luck after the golf task was not statistically significant, $t(109) = 1.78$, $p = .08$.

Could these results be due to insufficient superstition in our participants? This seems unlikely. Seventy percent of control and 80% of superstition-activated participants reported a belief in luck, similar to the target study’s participant pool. Moreover, excluding participants in both groups who did not believe in luck (using same criterion as Damisch, 2008) did not reveal an effect (see Supplementary Table S2, $t(82) = -0.69$, $p = .49$).

In exploratory analyses, we did not observe interactions by gender or experimenter, and excluding psychology majors did not have an effect (see Supplementary Table S2).

Aggregating the data across the two studies indicates a null effect of superstition on performance: Unbiased overall $d = 0.05$, 95% CI $[-0.21, 0.30]$, as indicated by the black diamond in Figure 1. The confidence interval of this estimate does not overlap with that generated across the studies by Damisch et al. (2010) and Damisch (2008) (white diamond, Figure 1).

Meta-Analysis

To better understand the divergence between our results and those of the target study, we conducted a small-scale meta-analysis. We included the original golf experiment and conceptual replications described in the introduction (Aruguete et al., 2012; Damisch, 2008; Damisch et al., 2010; Lee et al., 2011) and the two attempts reported here (summarized in Table 1 and Figure 1). We conducted the

meta-analysis using ESCI (Cumming, 2011), an Excel-based analysis package which includes tools for integrating effects sizes and visualizing their differences across studies.

The meta-analysis provides an overall unbiased estimate of effect size: $d = 0.40$, 95% CI $[0.14, 0.65]$, gray diamond in Figure 1. However, there is significant heterogeneity in the reported effect sizes ($Q(10) = 26.52$, $p = .003$): One subset of studies indicates a strong effect, the remainder indicates little to no effect. This heterogeneity requires caution in interpreting the overall estimated effect size (see Discussion).

General Discussion

Although we took care to precisely replicate the materials and procedures of the target study, we could not replicate the strong effect of superstition on performance consistently observed by Damisch et al. (2010) and Damisch (2008).

What could account for our failed replications? We can rule out a lack of impact: We observed robust effects in a manipulation check, conducted a second replication that achieved an even higher impact, and implemented quality controls that allowed filtering out of any participants not sufficiently engaged in the task. It is possible that the target study achieved even greater impact but no manipulation check was conducted to provide comparison. This seems implausible, however, as Damisch et al. (2010) was able to observe strong effects on performance with subtle manipulations.

Our meta-analysis suggests considerable heterogeneity in observed effects of superstition on performance. Such heterogeneity can indicate the operation of a moderator, perhaps one that differs between the European participants in the target study and the American participants in these replications. Indeed, culture can play a surprisingly large role even in basic psychological phenomena (Henrich, Heine, & Norenzayan, 2010). This seems unlikely, however, as we took care to equate with the original study or monitor key moderators including belief in luck, task difficulty, and sample characteristics. It is notable, though, that in the Damisch et al. studies (2008) performance gains in the superstition group were associated with increased self-efficacy (Study 3 and 4) and task persistence (Study 4). This suggests, then, that strong effects of superstition only emerge when control participants are not confident or motivated enough to perform near their ability, providing “room” for superstition to boost performance through these factors. Indeed, Matute (1994) and others have suggested that superstitions function specifically to maintain performance under adverse conditions.

Heterogeneity of effect sizes can also arise due to substantive differences in research quality. We made every effort to replicate the target study precisely. Further, we developed an *a priori* sampling plan, took steps to minimize expectation effects (e.g., experimental script), and acquired a large enough sample size to provide a relatively precise estimate of effect size. These are all design features recently emphasized for increasing research rigor,

especially for ensuring good control of Type I error (e.g., Simmons, Nelson, & Simonsohn, 2011). Along these lines, it is notable that the four studies with these features (our own plus the two from Aruguete et al., 2012) consistently indicate no effect of superstition on performance. The studies that do show an effect of superstition on performance lack some or all of these design features. Moreover, the Damisch studies show a remarkable consistency of result that could occur if Type I error is not well controlled: Given the overall effect size from these studies (0.83, white diamond, Figure 1), the odds of all six of these studies reaching statistical significance is only four in 100.

Ultimately, only further research can determine if the lack of effect we observed is due to moderators, improved rigor, or both. Currently, the studies with the strongest design features do not indicate a robust effect of superstition on performance.

Acknowledgments

We thank Jamie Mussen, Malgorzata Rozko, and Amy Taraszka for collecting the data for this study. We also thank Lysann Damisch for her gracious and extensive cooperation. Funding for materials and participation remuneration was provided by a grant from the Center for Open Science. The authors declare no conflict-of-interest with the content of this article. Designed Research: R. C-J., T. C.; Analyzed Data: R. C-J.; Wrote Paper: R. C-J., T. C. All materials, data, video of the procedure, and the preregistered design are available at <http://osf.io/fsadm/>.



References

- Aruguete, M. S., Goodboy, A. K., Jenkins, W. J., Mansson, D. H., & McCutcheon, L. E. (2012). Does religious faith improve test performance? *North American Journal of Psychology*, 14, 185–196.
- Cumming, G. (2011). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. New York, NY: Routledge.
- Damisch, L. (2008). *Keep your fingers crossed! The influence of superstition on subsequent task performance and its mediating mechanism* (Doctoral dissertation, University of Cologne, Cologne, Germany). University of Cologne. Retrieved from <http://kups.ub.uni-koeln.de/2471/>
- Damisch, L., Stoberock, B., & Mussweiler, T. (2010). Keep your fingers crossed! How superstition improves performance. *Psychological Science*, 21, 1014–1020. doi: 10.1177/0956797610372631
- Doheny, L. (2010, July 16). Good Luck Charms Might Just Work. *U.S. News Health*. Retrieved from <http://health.usnews.com/health-news/family-health/brain-and-behavior/articles/2010/07/16/good-luck-charms-might-just-work>
- Dupont, W. D., & Plummer, W. D. (1998). Power and sample size calculations for studies involving linear regression. *Controlled Clinical Trials*, 19, 589–601.
- Galton, F. J. (1872). Statistical inquiries into the efficacy of prayer. *The Fortnightly Review*, 12, 125–135. doi: 10.1093/ije/dys109
- Good-Luck-Gifts.com. (n.d.). *Lucky charms really work*. Retrieved from <http://www.good-luck-gifts.com/index.php/lucky-charms-guide/118-lucky-charms-really-work>
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *The Behavioral and Brain Sciences*, 33, 61–63. doi: 10.1017/S0140525X0999152X
- Hutson, M. (2010, February). *Harder, Better, Faster, Luckier [Web log post]*. *Psychology Today, Psyched! Blog*. Retrieved from <http://www.psychologytoday.com/blog/psyched/201002/harder-better-faster-luckier>
- Lee, C., Linkenauger, S. A., Bakdash, J. Z., Joy-Gaba, J. A., & Proffitt, D. R. (2011). Putting like a pro: The role of positive contagion in golf performance and perception. *PloS One*, 6, e26016. doi: 10.1371/journal.pone.0026016
- Matute, H. (1994). Learned helplessness and superstitious behavior as opposite effects of uncontrollable reinforcement in humans. *Learning and Motivation*, 25, 216–232. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0023969084710125>
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22, 1359–1366. doi: 10.1177/0956797611417632
- Skinner, B. (1948). “Superstition” in the pigeon. *Journal of Experimental Psychology*, 38, 168–172. Retrieved from <http://psycnet.apa.org/journals/xge/38/2/168/>

Received February 18, 2013
Accepted December 3, 2013
Published online May 19, 2014

Robert J. Calin-Jageman

Department of Psychology
Dominican University
7900W. Division
River Forest, IL 60305
USA
E-mail rcalinjageman@dom.edu