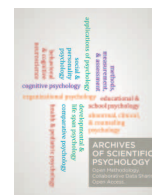




AMERICAN  
PSYCHOLOGICAL  
ASSOCIATION

## Archives of Scientific Psychology

<http://www.apa.org/pubs/journals/arc>



# Can Physical Warmth (or Coldness) Predict Trait Loneliness? A Replication of Bargh and Shalev (2012)

Jessica Wortman, M. Brent Donnellan, and Richard E. Lucas  
Michigan State University

### ABSTRACT

In recent years, psychologists have become increasingly interested in the extent to which abstract concepts can be “embodied” in physical experiences. Bargh and Shalev (2012) demonstrated that individuals who experienced physical coldness (in the form of a cold pack) reported that they were lonelier than individuals who experienced physical warmth (in the form of a hand warmer) (Bargh & Shalev, 2012). Using procedures that were very similar to those in the original study and a sample size that was more than 5 times larger, we found that there was no difference between conditions, a finding that failed to replicate the original study. People who held a cold pack did not report that they were lonelier than people who held a warm pack. Holding a cold pack versus a warm pack also did not have an effect on people’s personality traits. Overall, we suggest that there needs to be further research to determine if there is a connection between physical warmth and interpersonal warmth.

### SCIENTIFIC ABSTRACT

Bargh and Shalev (2012) hypothesized that experiencing physical coldness will lead individuals to report greater loneliness than if they experienced physical warmth. In their Study 2, they conducted an experiment in which they showed that participants who held a cold pack reported higher trait loneliness (as measured by a short form of the UCLA Loneliness Scale; Russell, 1996) than participants in the warm condition. We attempted to replicate this potentially practically important finding in a high-powered study ( $N = 260$ ). We also assessed the Big Five personality traits to determine if warmth or coldness might lead to changes in self-reported personality traits (particularly agreeableness). Our results showed that holding a hand warmer or cold pack for 1 min had no effect on trait loneliness in our study, with an effect size of essentially zero. The effect remained nonsignificant after excluding participants who reported any suspicion about the connection between the warmth-coldness manipulation and the measure of loneliness. There were also no effects of the cold (vs. warm) packs on personality traits. The question of the potential connection between physical warmth or coldness and loneliness warrants further research before it can be accepted.

**Keywords:** loneliness, replication, effect sizes

**Supplemental materials:** <http://dx.doi.org/10.1037/arc0000007.supp>

**Data Repository:** <http://dx.doi.org/10.3886/ICPSR35220.v1>

Can feelings of loneliness be induced by the experience of physical coldness? Recent psychological research suggests that the association between coldness and loneliness might transcend a purely linguistic

connection. In particular, Bargh and Shalev (2012) provided experimental evidence suggesting that participants who held cold packs reported higher scores on a measure of trait loneliness compared with participants who held warm packs. This research extended an earlier study by Williams and Bargh (2008) that suggested that individuals who held cold packs were less prosocial than individuals who held warm packs. These temperature priming findings received considerable attention in the scholarly community and in the popular press (e.g., Tierney, 2008). Such results may also prove to have practical value because they hint at the possibility that physical warmth might provide a way to alleviate mental health symptoms (Bargh & Shalev, 2012). However, there are concerns about the robustness of certain findings in this literature (e.g., Donnellan, Lucas, & Cesario, in press; Lynott et al., in press); thus, it is important to

This article was published August 18, 2014.

Jessica Wortman, M. Brent Donnellan, and Richard E. Lucas, Department of Psychology, Michigan State University.

For further discussion on this topic, please visit the *Archives of Scientific Psychology* online public forum at <http://arcblog.apa.org>.

Correspondence concerning this article should be addressed to Jessica Wortman, Department of Psychology, 316 Physics, Room 244B, Michigan State University, East Lansing, MI 48824. E-mail: [wortmanj@msu.edu](mailto:wortmanj@msu.edu)

conduct high-power replications of these kinds of findings. Accordingly, we conducted a replication of Study 2 in Bargh and Shalev (2012).

The original Bargh and Shalev experimental study included 75 participants (51% female) recruited outside of a university dining hall. Participants held either a cold or a warm pack for 1 min under the guise of evaluating a commercial product. Participants subsequently responded to three questions evaluating the product for its pleasantness, effectiveness, and whether they would recommend it to their friends (all on a *yes/no* response scale). The warm pack was heated in a microwave to a temperature of approximately 98°F whereas the cold pack was cooled in a freezer for 1 hr. There was also a control condition in which participants did not evaluate a product. Then, ostensibly as a part of another study, participants responded to a version of the UCLA Loneliness scale (see Russell, 1996). Results showed that holding a cold pack was associated with greater feelings of loneliness ( $M = 2.52$ ,  $SD = .91$ ) as compared with the warm pack condition ( $M = 2.04$ ,  $SD = .64$ ;  $d = .61$ , approximate 95% confidence interval [CI] on the  $d$  [.05, 1.19];  $p = .04$ ) and the no pack control condition ( $M = 1.97$ ,  $SD = .68$ ;  $d = .74$ ;  $p < .01$ , approximate 95% CI = [.17, 1.34]). However, there was no evidence for a difference between the warm pack and control conditions ( $M = 2.04$  vs.  $M = 1.97$ ;  $d = .11$ ,  $p > .05$ , approximate 95% CI = [-.45, .67]).

The theoretical idea underlying the connection between physical temperatures and loneliness draws upon recent research and theorizing in social cognition related to the idea of embodiment. The gist of embodiment research is that abstract concepts such as social closeness/loneliness are linked to physical concepts in the associative network that exists in the human mind. This associative network is presumably developed through early experiences that lead physical warmth or coldness to become closely associated with interpersonal connection or isolation (see Bargh & Shalev, 2012). One possible mechanism offered by Bargh and Shalev (2012) is that the protection and physical warmth provided by caregivers early in life becomes associated with psychological connectedness, a finding broadly compatible with Bowlby's (1969) attachment theory. Consistent with this proposition, there is emerging neurological research suggesting that brain systems involved in processing physical warmth overlap with systems processing social connection. For example, Inagaki and Eisenberger (2012) showed that participants' neural activity while reading a loving message (vs. a neutral message) overlapped with neural activity that occurred while holding a warm pack (vs. a ball). However, this imaging study was based on relatively few participants ( $N < 20$ ; see p. 2275).

The basic idea from the relevant embodiment literature is that experiences of physical coldness can produce psychological feelings of loneliness as demonstrated in Study 2 in Bargh and Shalev (2012). Moreover, the Bargh and Shalev (2012) Study 2 findings are consistent with a growing literature on the broader psychological effect of cold versus warm priming on interpersonal thoughts, feelings, and behaviors. For example, IJzerman and Semin (2009, 2010) demonstrated that various manipulations of physical warmth have an association with interpersonal warmth. Perceiving communal as opposed to agentic traits was associated with higher ratings of ambient temperature (IJzerman & Semin, 2010) whereas experiencing physical warmth (e.g., a warm room, a warm beverage) is associated with greater perceived closeness to an experimenter (IJzerman & Semin, 2009). Likewise, Williams and Bargh (2008) found that participants who held a warm pack were more prosocial than participants who held a cold pack.

Although there are positive cold versus warm priming effects documented in the literature, there are also failures to replicate results involving temperature and psychological outcomes. Donnellan et al.

(in press) attempted to duplicate results from Study 1a in Bargh and Shalev (2012) regarding a possible correlation between trait loneliness and a preference for warm showers/baths under the idea that lonely people actually substitute physical warmth for a lack of social warmth. Across nine studies involving over 3,000 participants, Donnellan et al. did not duplicate the original effect size estimates. They found that the overall correlation between warm water and trait loneliness was statistically indistinguishable from zero using a random effects meta-analysis that included the Bargh and Shalev (2012) data. This meta-analytic "null" result is consistent with subsequent failures to replicate reported by Donnellan et al. (In press) and Ferrell, Gosling, and Donnellan (2014). Lynott and colleagues (in press) attempted three replications of the Williams and Bargh (2008) study linking cold versus warm pack priming with prosocial behavior. Each replication attempt had 4 times the original sample size of 50 participants, and each study failed to duplicate the original result. Williams et al. (in press) summarized the effect size estimates for studies investigating the effect of temperature primes on prosocial outcomes and the associated sample sizes. They noted that the largest sample size was less than 60 participants. Literatures based on small samples are thought to have increased risks of Type I errors and to produce inflated effect size estimates (see Button et al., 2013; Ioannidis, 2005, 2008).

Given these concerns about the existing literature, it is important to conduct systematic investigations of specific effects so that researchers have a better understanding of the robustness of specific effects and a better understanding of the underlying effect sizes. Thus, we decided to test the robustness of Study 2 in Bargh and Shalev (2012) to extend previous correlational research by Donnellan et al. (in press). We attempted a near-exact replication of this experiment with some modifications to facilitate data collection from our university subject pool. Rather than using a therapeutic warm pack that was heated in a microwave, we chose instead to use instant hand warmers that would maintain a consistent temperature for several hours. Likewise, we used instant cold packs that were squeezed to activate for the sake of convenience and to provide a relatively consistent temperature across participants without the need of a freezer. We did not include a no-pack control condition because there was no evidence of a statistical difference between the warm pack and the no pack control condition in the original study. This change in procedure also helped bolster the cover story. Bargh and Shalev (2012) noted that the effect of cold manipulations tended to be more pronounced than the effect of warmth manipulations (p. 158); therefore, we believed this procedural change would also make for a more efficient replication attempt because we would only need two conditions as opposed to three. Otherwise, procedures were similar to the original study. We wanted to have at least .99 power to detect the original effect size estimate of .61 between the cold pack and warm pack conditions, which meant that we would need at least 100 participants in each condition (we would need only 44 in each condition if we wanted .80 power). We collected data from more than 200 participants in case we had to discard suspicious participants.

## Method

### Participants

Participants were 260 college students (77.7% women, with 95% of participants between 18 and 22) who completed the study for course credit at Michigan State University. Most participants self-identified as White (79.2%). Participants were randomly assigned to either the cold ( $n = 137$ ) or warm ( $n = 123$ ) pack condition. Data were

collected at the end of the fall semester in 2012 ( $n = 101$ ) and at the start of the spring semester in 2013 ( $n = 159$ ).

## Materials

Participants in the cold condition held a 4"  $\times$  5" instant cold pack (from First Aid store.com Vancouver, WA). The cold pack was activated when trained assistants squeezed the packs before providing it to participants in the cold condition. The packs remained cold (approximately 40 °F) for approximately 15 min after the initial activation. The activated cold packs were discarded after each experimental session (every 30 min). Participants in the warm condition held a Little Hotties brand hand warmer. Because the hand warmer stayed warm for up to 10 hr, research assistants were instructed to use the same hand warmer across experimental sessions to minimize waste. In between experimental sessions, the hand warmer was wrapped in a cloth to maintain its temperature at approximately 100 °F. The assistants tested the hand warmer before each session to ensure it was warm before giving it to participants. In all conditions, participants set the pack on a single layer of tissue in their palm (because the packs were not intended to be applied to bare skin for safety reasons), set another layer of tissue on top, and rested their other hand on top.

## Procedure

Participants came into the laboratory (up to three at a time). They were informed that they would be participating in two unrelated studies that were combined into a single laboratory session because both studies were short. Students were awarded one experimental credit per 30 min of participation in research so that the cover story was consistent with this procedure at our institution. Participants were told they would first be evaluating a product and answering some brief questions. They were then told they would then participate in a short pilot study to evaluate personality items for an inventory being created by researchers at their institution. Participants completed two separate informed consent forms to emphasize that the studies were unrelated.

After providing consent, the research assistant (who was blind to the hypothesis) led each participant to separate rooms where they completed the tasks on a computer. Research assistants entered a random identification number for each participant (to preserve their anonymity) and then the computer randomly assigned that participant to test either a warm or a cold pack. This kept research assistants blind to condition for as long as possible. We did not believe it would be feasible to keep assistants completely blind because we wanted to monitor how long participants held the packs and to make sure that no participant experienced discomfort. Researchers in the original study were not blind to condition.

Research assistants retrieved the correct pack for the participant from a separate room and activated the cold pack when required. The participant was instructed to hold the pack (in a layer of tissue) in their hand for 1 min, and the research assistant used a stopwatch to time exposure. Participants were instructed to stop holding the pack if it became too uncomfortable or painful. No participant stopped holding a pack early. Once the participant had held the pack for 1 min, the research assistant instructed them to complete the surveys. Participants were then left alone. Research assistants then assisted other participants or went to a separate room and waited for the experiment to end.

Participants first responded to questions intended to evaluate the effectiveness of the products (see Appendix A). There was a brief break between the two surveys informing the participants when they had completed Study 1 (the product evaluation study) and would be

moving on to Study 2. Participants then rated the items to assess chronic or trait loneliness as well as a short 20-item Big Five personality measure (Donnellan, Oswald, Baird, & Lucas, 2006; see Appendixes B and C). The Big Five measure was included to reduce suspicion by having more items in the personality survey. This addition also allowed us to evaluate an exploratory hypothesis as to whether exposure to a warm or cold pack would affect responses to personality measures especially with respect to agreeableness, the Big Five trait related to prosocial behavior. It is important to note that the personality trait measure was administered after the loneliness measure so that the procedures were as consistent as possible with the original study.

Finally, participants answered a closed-ended question about the temperature of the pack as a check on manipulation and then answered a series of questions designed that allowed for a funnel debriefing procedure. These were also completed on the computer. A complete copy of the items is in Appendix D.

## Measures

**Product evaluation items.** Participants rated the pack on whether it was effective (1 = *yes* 2 = *no*), pleasant (*yes/no*), and whether they would recommend the product to their friends (*yes/no*). These were based on the descriptions provided in Bargh and Shalev (2012).

**Personality items.** Participants first completed the 10-item loneliness scale used in Bargh and Shalev (2012). This measure was a modified form of the first version of the UCLA Loneliness scale (Russell, Peplau, & Ferguson, 1978). Donnellan et al. (in press) reported that scores on this measure correlated .82 with the full 20-item UCLA Loneliness scale (Russell, 1996), a scale commonly used in contemporary loneliness research. Participants responded on a 1 (*never*) to 4 (*often*) scale indicating the frequency of feeling loneliness (e.g., "How often do you feel completely alone?"). The measure was scored so that higher scores indicating greater loneliness ( $M = 2.10$ ,  $SD = .61$ , Cronbach's  $\alpha = .89$ ). Participants completed the 20-item mini-International Personality Item Pool (IPIP; Donnellan et al., 2006; see also Cooper, Smillie, & Corr, 2010), which assesses the Big Five personality domains with four items per each domain. Participants rated how accurately the statement described them on a 1 (*very inaccurate*) to 5 (*very accurate*) Likert scale. Descriptive statistics and reliabilities for this measure are presented in Table 1.

**Manipulation and suspicion check items.** Participants rated the temperature of the pack on a 1 (*very cold*) to 6 (*very hot*) scale as a manipulation check. Then, participants responded to open-ended questions, reporting whether they thought anything about the study was strange, as well as their likes and dislikes about the study, anything they thought that we should change, and general comments. Finally, participants indicated whether or not they believed the purpose of the study ("Did you believe the purpose of the study that we told you?") on a *yes/no/maybe* scale.

## Results

### Preliminary Analyses

Warm packs were rated more favorably than cold packs, although participants had generally positive evaluations of both packs. In terms of pleasantness, 95.9% of the participants in the warm condition answered yes to this question ( $n = 118$  out of 123) versus 79.6% in the cold condition ( $n = 109$  out of 137) ( $\chi^2 = 15.68$ ,  $df = 1$ ,  $p < .05$ ). In terms of effectiveness, 98.4% of the participants in the warm condition answered yes to this question ( $n = 121$  of 123) versus 87.6% in the cold condition ( $n = 120$  of 137) ( $\chi^2 = 11.125$ ,  $df = 1$ ,



Table 1  
Descriptive Statistics and Hypothesis Tests for the Mini-IPIP

	Cronbach's $\alpha$	Overall $M$ ( $SD$ )	Cold Condition $M$ ( $SD$ )	Warm Condition $M$ ( $SD$ )	Warm-Cold $t$	Warm-Cold $p$	Warm-Cold Cohen's $d$
Extraversion	.83	3.25 (.94)	3.32 (.97)	3.18 (.90)	1.15	.25	.15
Agreeableness	.71	4.09 (.65)	4.10 (.66)	4.09 (.65)	0.08	.93	.01
Conscientiousness	.71	3.58 (.74)	3.65 (.76)	3.50 (.71)	1.55	.12	.19
Neuroticism	.70	2.73 (.81)	2.73 (.84)	2.74 (.78)	-0.18	.86	-.02
Openness	.78	3.66 (.79)	3.65 (.81)	3.68 (.78)	-0.30	.76	-.04

Note. Warm-Cold  $t$  refers to the independent group  $t$  test comparing the warm and cold pack conditions on the Big Five traits. For all  $t$  tests presented here, the  $df = 258$ . Positive effect sizes indicate that the cold group scored higher than the warm group.

$p < .05$ ). Lastly, 95.1% of participants in the warm condition indicated they would recommend the pack to friends (117 of 123) versus 84.7% of participants in the cold condition (116 of 137) ( $\chi^2 = 7.606$ ,  $df = 1$ ,  $p < .05$ ). Most critically for the manipulation check, the packs were rated differently in terms of subjective rating of temperature ( $t = 34.391$ ,  $df = 189.206$  using a test for unequal variances). Cold packs were rated as colder than warm packs ( $M$  cold = 1.93,  $SD = .43$  vs.  $M$  hot = 4.57,  $SD = .75$ ). Given that the variances were unequal, it is difficult to justify a standardized effect size estimate. However, this is an apparently substantial difference that appears consistent with similar work that has manipulated hot and cold using instant packs when considering the effect on subjective ratings of exact temperature (i.e.,  $t = 10.409$ ,  $df = 46$ ,  $d = 3.07$ , 95% CI = [2.28, 3.99] in Williams & Bargh, 2008).

### Critical Analyses

An independent sample  $t$  test indicated that there was no average difference in loneliness between participants in the warm ( $M = 2.10$ ,  $SD = .60$ ) versus cold ( $M = 2.11$ ,  $SD = .62$ ) conditions,  $t(258) = -0.006$ ,  $p = .995$ . The effect size for this difference was almost exactly zero— $d = .02$  (95% CI on the  $d$  of [-.23, .26]; scored so that positive scores indicate the cold pack group was higher on the trait loneliness measure). Thus, we did not duplicate the original results either in terms of effect sizes or statistical significance. Recall that the original effect size estimate was approximately .61 but had a large approximate 95% CI of between .05 and 1.19. Therefore, we tried to compare the effect size estimate we obtained with that of the original study to determine if we were able to statistically reject the hypothesis that our respective effect size estimates came from the same population. We converted the  $d$ s to  $r$ s and used the test for the difference between independent correlations for our significance test. Results suggested that our effect size estimate is not statistically different from that of the original study ( $z = 1.85$ ,  $p = .064$ ), although we have a narrower CI. Asendorpf et al. (2013, p. 112) noted that formal comparisons of effect sizes can be uninformative when trying to compare replication results to initial studies based on small samples because of the width of the initial CI.

### Supplemental Analyses

The correlation between subjective ratings of temperature and loneliness was not statistically significant ( $r = .002$ ,  $p = .976$ ,  $n = 259$ ); thus, there was no evidence for an association between temperature and loneliness across the sample. The correlation between loneliness and a composite of the pack evaluation items, scored such that higher values represent more positive evaluations, was also not statistically significant ( $r = -.054$ ,  $p = .383$ ,  $n = 260$ ). This means there was little evidence of an association between loneliness and how the packs were evaluated by participants. Likewise, there was no statistical

effect of condition controlling for positive evaluations of the pack in a model with both as predictors of loneliness (standardized regression  $\beta = -.017$ ,  $p = .790$ ). This suggests that there is no linear association between the condition and loneliness controlling for the positive evaluation of the pack.

We also conducted a series of sensitivity checks to determine whether analyzing different subsets of participants would alter our conclusions. We first conducted separate  $t$  tests for the participants in the fall and spring, and both tests yielded null results (fall:  $t(99) = -0.334$ ,  $p = .739$ ,  $d = .07$ , approximate 95% CI = [-.33, .43], cold group higher; spring:  $t(157) = 0.239$ ,  $p = .811$ ;  $d = -.04$ , approximate 95% CI = [-.35, .27], cold group lower). There was no evidence of a mean difference between semesters in terms of loneliness ( $t(258) = 0.104$ ,  $p = .917$ ;  $d = -.02$ , approximate 95% CI = [-.33, .30], spring group lower). We then noted that some of the participants responded to the manipulation checks in ways that may undermine their experiences of the experimental conditions. In particular, there were participants in the warm condition that reported that the pack was lukewarm ( $n = 7$ ), and participants in the cold condition that reported the pack was either lukewarm ( $n = 5$ ) or warm ( $n = 1$ ). However, the results remained virtually unchanged when these participants were excluded ( $t(245) = .17$ ,  $p = .87$ ,  $d = -.02$ , approximate 95% CI = [-.27, .23], cold group lower and the effect size remained essentially zero).

We also tested whether participants' suspicion might have undermined our ability to detect the experimental effect. We asked participants whether they believed the purpose of the study and the modal response was "maybe" believed the purpose of the study ( $n = 113$ ), with "yes" being the next most frequent response ( $n = 82$ ), followed by "no" ( $n = 63$ ). Three participants did not respond to this question. We began by only analyzing those that answered affirmatively to this question ( $n = 82$ ). Participants in the warm condition ( $M = 2.13$ ,  $SD = .64$ ) did not report lower levels of loneliness than those in the cold condition ( $M = 2.02$ ,  $SD = .64$ ,  $t(80) = -.74$ ,  $p = .46$ ,  $d = -.17$ , approximate 95% CI = [-.61, .27], cold group lower using this restricted set). In fact, the difference was in the opposite direction as predicted, and this sample size was still larger than the original study. Moreover, restricting analyses to these participants did not change any of the findings concerning effectiveness, pleasantness, or likelihood of recommending the pack. We also failed to find a difference when we just excluded those who answered no to this question. Participants in the warm condition ( $M = 2.10$ ,  $SD = .60$ ) reported slightly higher levels of loneliness than those in the cold condition ( $M = 2.04$ ,  $SD = .62$ ,  $t(193) = -.67$ ,  $p = .51$ ,  $d = -.10$ , approximate 95% CI = [-.38, .18]).

We then examined open-ended responses to the debriefing questions to determine if excluding participants who made a connection between the two studies might have an effect on the results. We used a very strict decision rule, excluding any participants who indicated

that they believed the two studies might be related. Approximately 50% ( $n = 136$ ) indicated that they believed that there was some connection between the two studies. Thus, we excluded those individuals, leaving a sample size of 124. However, we still found no indications of differences in loneliness between the cold condition ( $M = 1.97$ ,  $SD = .56$ ) and the warm condition ( $M = 2.08$ ,  $SD = .56$ ,  $t(122) = 1.04$ ,  $p = .30$ ,  $d = .19$ , approximate 95% CI =  $[-.17, .55]$ ). In fact, the observed difference was actually counter to the Bargh and Shalev (2012) predictions because the cold pack group reported lower average scores. It is worth noting that only two participants correctly identified the purpose of the study—one indicating that warmth might make you feel less lonely and the other suggesting that warmth might make you feel more “connected” to others. Most of the open-ended responses indicated that participants believed that we were testing the connection between product evaluations and individuals’ personalities.

### Exploratory Big Five Results

Last, we tested to see if the manipulation might have any effect on participants’ self-ratings of personality. In particular, we wanted to examine if there might be differences in self-ratings of agreeableness given prior work showing that warm (vs. cold) packs lead to more prosocial behavior (Williams & Bargh, 2008). The results of these  $t$  tests, along with the means and standard deviations for each of the conditions, are presented in Table 1. We found no evidence of statistically significant differences in any of the Big Five personality traits between the two conditions. There was a nonsignificant trend for participants in the cold condition to rate themselves higher on conscientiousness than participants in the warm condition. However, the difference was still quite small ( $d = .19$ ) and would not have been predicted by theory. Exposure to a hot or cold pack seems to have had no significant effect on participants’ average self-ratings of personality traits.

### Discussion

This study was a near-exact replication of Study 2 as reported in Bargh and Shalev (2012) with a sample size that was approximately 5 times larger than the combined sample size from the two relevant conditions from the original study. We did not find any differences in trait loneliness between participants in the cold condition and those in the warm condition. Our effect size estimates were close to zero and were more precise than the original effect size estimate of .61 from the original study. In light of these results, we suggest that more work is needed to determine whether there is compelling evidence that differences in trait loneliness can be induced after holding a cold or warm pack after 1 min and to increase precision in the effect size estimate.

Of course, it is important to acknowledge that there are many potential explanations for the discrepancy between these findings and those reported in Bargh and Shalev (2012). Some authors have suggested that priming effects might be sensitive to variations in experimental procedures or theoretically relevant moderators, causing them to be difficult to detect (e.g., Cesario, 2014). Bargh and Shalev (2012) did not discuss potential moderators of these effects or propose expected limitations on the generalizability of their findings (see Donnellan et al., in press), but unexpected moderators are always a concern when replication attempts are unsuccessful. One possibility based on more recent research is that the attachment styles of participants may influence their susceptibility to cold versus warm priming effects (IJzerman, Karremans, Thomsen, & Schuber, 2013). IJzerman et al. (2013) found that securely attached children were more prosocial in a relatively warm classroom as opposed to a relatively cold classroom ( $d = .71$ ), a finding consistent with Williams

and Bargh (2008). However, there was no effect of classroom temperature on prosocial behavior for insecurely attached children. The suggestion then is that temperature priming effects are more likely to be found in samples of securely attached participants. Thus, future studies should test whether adult attachment variables moderate the effect of cold packs on loneliness in subsequent experiments.

It is also possible that subtle differences in the procedures between the original study and our own (i.e., a potential hidden procedural moderator) might explain our null results. Although we replicated their study as closely as we were able, we did so in a laboratory setting using a computer to collect survey responses and with a sample drawn from different student populations (Michigan State University students vs. Yale students). In addition, participants came to the laboratory in small groups but were in separate rooms during the experiment. Nonetheless, differences in the social context of the experimental setting might have been a factor. Finally, we used a slightly different manipulation of cold (vs. hot) here, relying on instant cold packs and hand warmers rather than heated or cooled therapeutic packs. Although we conducted manipulation checks to ensure that the packs used here were perceived as warm or cold, we acknowledge that this slight difference might have had some unintended consequence. Williams and Bargh (2008) used instant cold packs so they have been used in the literature, but it is still possible that pack differences may explain discrepancies.

We should also comment that the possibility of moderator effects would have implications for the real-world effect of the association between cold and loneliness. These potential moderators would suggest that experimental effects are contingent upon several seemingly auxiliary factors. If the ability to detect the effects of warm versus cold pack is dependent upon these kinds of factors, then there are limitations on the generalizability of the effects. This would undermine claims about the practical importance of the cold/warm priming effects on trait loneliness made in Bargh and Shalev (2012) because of the effect of uncontrollable moderators in the real world. Moreover, the current null findings are consistent with null results reported by Lynott et al. (in press) regarding the general inability of instant cold versus warm packs to generate detectable differences in prosocial behaviors. Moreover, the current findings are only relevant to the effect under investigation in Study 2 from Bargh and Shalev (2012) regarding the effect of cold packs on trait loneliness. To be sure, we emphasize that other studies have shown links between physical warmth and coldness in other domains such as estimation of the temperature of a room or the warmth of an experimenter (IJzerman & Semin, 2009, 2010).

All told, the link between exposure to a cold pack and loneliness needs further study to obtain a precise understanding of the population effect size and to determine whether this population effect size is meaningfully different from zero in a reasonably generalizable way. These results are not definitive by any means because they simply suggest that a degree of caution is warranted before assuming that the findings from Study 2 in Bargh and Shalev (2012) are easily replicable. Likewise, additional work is needed to identify potential psychological moderators of these effects such as attachment variables (IJzerman et al., 2013) and methodological moderators such as the type of cold pack and differences in student populations. Regardless, future studies should use even larger sample sizes to study this particular effect. However, the current study suggests that researchers might want to exercise some caution when citing the original findings because the evidence in support of the idea that holding cold objects increases trait loneliness might not be clear cut.

### References

- Asendorpf, J. B., Conner, M., De Fruyt, F., De Houwer, J., Denissen, J. J. A., Fielder, K., . . . Wicherts, J. M. (2013). Recommendations for increasing

- replicability in psychology. *European Journal of Personality*, 27, 108–119. doi:10.1002/per.1919
- Bargh, J. A., & Shalev, I. (2012). The substitutability of physical and social warmth in daily life. *Emotion*, 12, 154–162. doi:10.1037/a0023527
- Bowlby, J. (1969). Attachment and loss: Vol. 1. *Attachment*. New York, NY: Basic Books.
- Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, 14, 365–376. doi:10.1038/nrn3475
- Cesario, J. (2014). Priming, replication, and the hardest science. *Perspectives on Psychological Science*, 9, 40–48. doi:10.1177/1745691613513470
- Cooper, A. J., Smillie, L. D., & Corr, P. J. (2010). A confirmatory factor analysis of the Mini-IPIP five-factor model personality scale. *Personality and Individual Differences*, 48, 688–691. doi:10.1016/j.paid.2010.01.004
- Donnellan, M. B., Lucas, R. E., & Cesario, J. (in press). On the association between loneliness and bathing habits: Nine replications of Bargh and Shalev (2012). Study 1. *Emotion*.
- Donnellan, M. B., Oswald, F. L., Baird, B. M., & Lucas, R. E. (2006). The mini-IPIP scales: Tiny-yet-effective measures of the Big Five factors of personality. *Psychological Assessment*, 18, 192–203. doi:10.1037/1040-3590.18.2.192
- Ferrell, J. D., Gosling, S. D., & Donnellan, M. B. (2014, April 30). Showering and Loneliness: U of T Replication. Retrieved 10:59, July 14, 2014 from <http://www.PsychFileDrawer.org/replication.php?attempt=MTg3>
- Ijzerman, H., & Semin, G. R. (2009). The thermometer of social relations: Mapping social proximity on temperature. *Psychological Science*, 20, 1214–1220. doi:10.1111/j.1467-9280.2009.02434.x
- Ijzerman, H., & Semin, G. R. (2010). Temperature perceptions as a ground for social proximity. *Journal of Experimental Social Psychology*, 46, 867–873. doi:10.1016/j.jesp.2010.07.015
- Ijzerman, H., Karremans, J. C., Thomsen, L., & Schubert, T. W. (2013). Caring for sharing. *Social Psychology*, 44, 160–166.
- Inagaki, T. K., & Eisenberger, N. I. (2012). Neural correlates of giving support to a loved one. *Psychosomatic Medicine*, 74, 3–7. doi:10.1097/PSY.0b013e3182359335
- Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS Medicine*, 2, e124. doi:10.1371/journal.pmed.0020124
- Ioannidis, J. P. (2008). Why most discovered true associations are inflated. *Epidemiology*, 19, 640–648. doi:10.1097/EDE.0b013e31818131e7
- Lynott, D., Corker, K. S., Wortman, J., Connell, L., Donnellan, M. B., Lucas, R. E., & O'Brien, K. (in press). Replication of “Experiencing physical warmth promotes interpersonal warmth” by Williams & Bargh (2008, Science). *Social Psychology*.
- Russell, D. W. (1996). UCLA Loneliness Scale (Version 3): Reliability, validity, and factor structure. *Journal of Personality Assessment*, 66, 20–40. doi:10.1207/s15327752jpa6601\_2
- Russell, D., Peplau, L. A., & Ferguson, M. L. (1978). Developing a measure of loneliness. *Journal of Personality Assessment*, 42, 290–294. doi:10.1207/s15327752jpa4203\_11
- Tierney, J. (2008, October 23). Heart-warming news on hot coffee. *New York Times*. Retrieved from <http://tierneylab.blogs.nytimes.com/2008/10/23/heart-warming-news-on-coffee/?ei=5070&emc=eta1>
- Williams, L. E., & Bargh, J. A. (2008). Experiencing physical warmth promotes interpersonal warmth. *Science*, 322, 606–607. doi:10.1126/science.1162548
- Williams, L. E., Corker, K. S., Lynott, D., Wortman, J., Connell, L., Donnellan, M. B., . . . O'Brien, K. (in press). Commentary and rejoinder on Lynott et al. (2014). *Social Psychology*.

## Appendix A

### Product Evaluation Items

- |  |   |
|--|---|
| <p>1. Was the application of the pack pleasant?</p> <p>1. Yes</p> <p>2. No</p> <p>2. Was the application of the pack effective?</p> <p>1. Yes</p> <p>2. No</p> | <p>3. Would you recommend this pack to your friends?</p> <p>1. Yes</p> <p>2. No</p> |
|--|---|

## Appendix B

### Bargh and Shalev (2012) Loneliness Scale

Please indicate how often each of the statements below is descriptive of you. Please circle one number for each statement:

1 indicates “I *never* feel this way”

2 indicates “I *rarely* feel this way”

3 indicates “I *sometimes* feel this way”

4 indicates “I *often* feel this way”

1. How often do you feel unhappy doing so many things alone?
2. How often do you feel you have nobody to talk to?
3. How often do you feel you cannot tolerate being so alone?
4. How often do you feel as if nobody really understands you?

5. How often do you find yourself waiting for people to call or write?

6. How often do you feel completely alone?

7. How often do you feel you are unable to reach out and communicate with those around you?

8. How often do you feel starved for company?

9. How often do you feel it is difficult for you to make friends?

10. How often do you feel shut out and excluded by others?

*Note:* All items are exactly as reported in Bargh and Shalev (2012) and modified from Russell et al. (1978).

(Appendices continue)

## Appendix C

### Mini-IPIP-20

Use the rating scale below to describe how accurately each statement describes you. Describe yourself as you generally are now and not as you wish to be in the future. Please describe yourself honestly.

**1 = Very Inaccurate**

**2 = Moderately Inaccurate**

**3 = Neither Inaccurate nor Accurate**

**4 = Moderately Accurate**

**5 = Very Accurate**

1. Am the life of the party (E)
2. Sympathize with others' feelings (A)
3. Get chores done right away (C)
4. Have frequent mood swings (N)
5. Have a vivid imagination (O)
6. Don't talk a lot (E)
7. Am not interested in other people's problems (A)
8. Often forget to put things back in their proper place (C)

9. Am relaxed most of the time (N)
10. Am not interested in abstract ideas (O)
11. Talk to a lot of different people at parties (E)
12. Feel others' emotions (A)
13. Like order (C)
14. Get upset easily (N)
15. Have difficulty understanding abstract ideas (O)
16. Keep in the background (E)
17. Am not really interested in others (A)
18. Make a mess of things (C)
19. Seldom feel blue (N)
20. Do not have a good imagination (O)

*Note:* Items 6, 7, 8, 9, 10, 15, 16, 17, 18, 19, and 20 are reverse scored. E = Extraversion; A = Agreeableness; C = Conscientiousness; N = Neuroticism; and O = Openness/Intellect (see Donnellan et al., 2006).

## Appendix D

### Demographics and Funnel Debriefing Questionnaire

1. What is your gender?
  1. Male
  2. Female
  3. Other
2. What is your age in years?
  1. 18
  2. 19
  3. 20
  4. 21
  5. 22
  6. 23+
3. Are you Hispanic or Latino/a?
  1. Yes
  2. No
4. Please indicate your racial category
  1. American Indian/Alaskan Native
  2. Asian
  3. Native Hawaiian or Other Pacific Islander
  4. Black or African American
  5. White
  6. Other
5. Please mark "Rarely" for quality-control purposes
  1. Always
  2. Most of the time
  3. Sometimes
  4. Rarely

5. Never

Now we have some questions about the study. Answers to these questions will help us improve the study for the future. Thank you for your honest feedback.

1. First, what was the temperature of the pack you held?
  1. Very cold
  2. Cold
  3. Lukewarm
  4. Warm
  5. Hot
  6. Very hot
  7. I did not hold a pack.
2. Was there anything strange about this study? (Open-Ended)
3. Was there anything you liked about this study? (Open-Ended)
4. Was there anything you did not like about this study? (Open-Ended)
5. What should we change about this study? (Open-Ended)
6. What do you think was the point of this study? (Open-Ended)
7. Do you have any general comments? (Open-Ended)
8. Did you believe the purpose of the study we told you?
  1. Yes
  2. Maybe
  3. No

Received February 19, 2014  
 Revision received May 9, 2014  
 Accepted May 19, 2014 ■