Available online at www.sciencedirect.com
ScienceDirect
IJRM xx (2015) xxx–xxx
www.elsevier.com/locate/ijresmar

Replication Article

Reflections on the replication corner: In praise of conceptual replications☆,☆☆

John G. Lynch, Jr. [a,*], Eric T. Bradlow [b], Joel C. Huber [c], Donald R. Lehmann [d]

[a] Leeds School of Business, University of Colorado-Boulder, Boulder CO 80309, USA
[b] Wharton School, University of Pennsylvania, Philadelphia, PA 19104, USA
[c] Fuqua School of Business, Duke University, Durham, NC 27708, USA
[d] Columbia Business School, Columbia University, New York, NY 10027, USA

Abstract

We contrast the philosophy guiding the Replication Corner at IJRM with replication efforts in psychology. Psychology has promoted "exact" or "direct" replications, reflecting an interest in statistical conclusion validity of the original findings. Implicitly, this philosophy treats non-replication as evidence that the original finding is not "real" — a conclusion that we believe is unwarranted. In contrast, we have encouraged "conceptual replications" (replicating at the construct level but with different operationalization) and "replications with extensions", reflecting our interest in providing evidence on the external validity and generalizability of published findings. In particular, our belief is that this replication philosophy allows for both replication and the creation of new knowledge. We express our views about why we believe our approach is more constructive, and describe lessons learned in the three years we have been involved in editing the IJRM Replication Corner. Of our thirty published conceptual replications, most found results replicating the original findings, sometimes identifying moderators.

Keywords: Conceptual replication; Replication and extension; Direct replication; External validity

1. Introduction

A number of researchers have concluded that there is a replication crisis in the social sciences, raising new questions about the trustworthiness of our evidence base (e.g. Pashler & Wagenmakers, 2012; Simons, 2014). This issue is playing out in the academic psychology literature via a large-scale effort to conduct "direct" or "exact" replications of important papers. The findings are mixed, which has led to considerable acrimony and suspicion about the "replication police" (Gilbert, 2014) and "negative psychology" (Coan, 2014) with public shaming of authors whose work is found not to replicate. This is not true of all direct replication efforts: the just-released report by the Open Science Collaboration (2015) is a model of circumspection. After summarizing attempts to replicate 100 papers sampled from the 2008 issues of three top psychology journals, the authors note, "It is also too easy to conclude that a failure to replicate a result means that the original evidence was a false positive" (p. aac4716-6).

We believe that the Replication Corner in IJRM provides an alternative and perhaps more constructive model to direct replication efforts. It has encouraged papers that are either "replications and extensions" or "conceptual replications."

☆ The authors thank Joe Harvey and Andrew Long for their assistance with the coding of published and in press Replication Corner articles that appears in Tables 1 and 2, and they thank the authors of the 30 replications for providing feedback on both tables. They thank Carl Mela and Uri Simonsohn for comments on earlier drafts. They thank editors Jacob Goldenberg and Eitan Muller for comments and for creating the Replication Corner.
☆☆ Their support and vision has been invaluable. Finally, they thank IJRM Managing Editor, Cecilia Nalgon, for support of the replication corner and for preparing materials for our coding of articles.

* Corresponding author at: Ted Anderson Professor of Free Enterprise, Leeds School of Business, University of Colorado-Boulder, Boulder CO 80309.
E-mail addresses: john.g.lynch@colorado.edu (J.G. Lynch), ebradlow@wharton.upenn.edu (E.T. Bradlow), joel.huber@duke.edu (J.C. Huber), drl2@coloumbia.edu (D.R. Lehmann).

http://dx.doi.org/10.1016/j.ijresmar.2015.09.006
0167-8116© 2015 Published by Elsevier B.V.

Conceptual replications attempt to replicate an important original finding while acknowledging differences in background factors (e.g., participants, details of procedures) compared with the original article. Put differently, they attempt to study the same construct-to-construct relations as the original article despite operationalizing those constructs differently. Replications and extensions test both the original finding and a moderator of it. They may somewhat closely replicate the original in a part of the design but then test whether varying some specific background factor moderates the results.

We believe that the very concept of an "exact replication" in social science is flawed. Even if one used the exact same procedures, respondents may have changed over time. Exact replication is impossible. Therefore, the only issue is how close the replication is to the original, and whether it is desirable to be "as close as possible." When the goal is generalization, we argue that "imperfect" conceptual replications that stretch the domain of the research may be more useful.

The other problem with some current replication efforts is the focus by many on consistency of the original and the replicate in statistical significance of the effect as opposed to its effect size. We see scholars attempting "direct" replications who declare a replication failure if the original study found a significant effect and the replicate did not when they have neither shown that the effect sizes for replicate and original differ significantly nor that the pooled effect size is indistinguishable from zero. Maniadis, Tufano, and List (2014) declare a failure to replicate Ariely, Loewenstein, and Prelec's (2003) findings of anchoring effects on valuations of goods and experiences. But Simonsohn, Simmons, and Nelson (2014) show that the effect sizes in the replicate and original were entirely consistent. In particular, the confidence-interval around the estimate of the replication included the entire confidence-interval for the original study: it was simply the case that the replication had lower power due to lower sample sizes and floor effects. Because the point estimate effect size from Maniadis et al. was lower than that from Ariely et al., from a meta-analytic perspective, the results of Maniadis et al. properly should decrease one's estimate of the population effect size from Ariely et al. But the new results should strengthen belief that the effect in Ariely et al. is larger than 0 in the population — not decrease the belief as Maniadis et al. conclude.[1] Even if the meta-analytic effect sizes shrink towards zero with the addition of a new study, the standard error may decline even more making one even more convinced about the existence of the effect.

We have therefore created policies in the Replication Corner to provide an outlet for conceptual replications and replications with extension rather than "direct" replications. Moreover, we have encouraged authors to take a meta-analytic view of how their results increase or decrease the strength of evidence of the original findings.

In this article, we give our views of what we have learned so far, and we articulate the philosophy that has guided our efforts at IJRM, contrasting our philosophy of conceptual replication with what seems to prevail in psychology's efforts to encourage exact replications.

2. The history of the replication corner

Trust in behavioral research was badly shaken when internationally famous professors Stapel, Sanna, and Smeesters all resigned their faculty positions after investigations uncovered unethical data reporting and outright fabrication of data. Similar misconduct was found in many other branches of science (e.g. Fang, Steen, & Casadevall, 2012).

Around the same time, there were highly publicized failures to replicate important findings even though no fraud was suspected, with contentious exchanges about the adequacy of the replication efforts. For example, Bargh, Chen, and Burrows (1996) found that priming elderly stereotypes caused subjects to walk more slowly and replicated the finding in their own paper. But Doyen, Klein, Pichon, and Cleeremans (2012) could not replicate the finding unless the experimenters were aware of the hypotheses being tested. Given growing questions about the reliability of published research, the journal *Perspectives on Psychological Science* then published a special section on replicability (Pashler and Wagenmakers (2012)).

An influential article by Simmons, Nelson, and Simonsohn (2011) pointed out that even researchers who were not trying to deceive their audiences might be deceiving themselves by ("field-allowed") flexibility in data analysis and reporting. They outlined common practices in data analysis that could inflate the effective type I error rate far in excess of the nominal 0.05 level, leading authors to find evidence of an effect that doesn't exist in the population. More generally, Gelman and Loken's (2014) analysis identified a broader "statistical crisis in science" resulting from overly flexible data analysis decisions.

In response to these disturbing revelations, the top journals in marketing and many in psychology have required more complete reporting of instruments and methods, greater specifics of data analysis, and more complete publication of materials. The goal was to allow scholars to better assess what is in the paper and to go beyond it to uncover further insights. For IJRM, Jacob Goldenberg and Eitan Muller took the view that replication must be an important complement to these disclosure requirements if we are to understand the scope and limits of our published findings. They noted that no high-status marketing journal was publishing successful or unsuccessful replication attempts and decided that IJRM would provide an appropriate outlet.

The approach of Jacob, Eitan and the Replication Corner co-editor team in 2012 differed from what was emerging in psychology. Below, we lay out here the difference in philosophy and take stock of what we have learned over the past three years.

[1] Simonsohn et al. point out that not only did Maniadis et al. successfully replicate Ariely et al. (2003): ironically, much of their paper presented a theoretical model of replications that was an exact replication of the model presented by Ioannidis (2005) in his widely cited "Why Most Published Research Findings are False."

Please cite this article as: Lynch, J.G., et al., Reflections on the replication corner: In praise of conceptual replications, *IJRM* (2015), http://dx.doi.org/10.1016/j.ijresmar.2015.09.006

- First, we decided to focus the replication corner on "important" and highly cited papers.
- Second, we expressed a preference for "conceptual replications" and "replications with extensions" rather than "direct" or "exact" replications. Put differently, our focus has been on matters of external validity rather than of the statistical conclusion validity of the original finding (cf. Cook & Campbell, 1979).
- Third, we have tried to be even-handed in publishing both "successful" replications of original findings and "failures to replicate."

3. Focus on "important" papers

We have chosen to consider only replication attempts of "important" papers. With limited pages, we wished to focus limited scholarly resources where it would matter most. Since we started the Replication Corner, psychology journals have followed a similar path.

A (much) secondary motive for our focus on "important" papers was to provide a further incentive for good professional behavior. We assume that most scholars in the field are earnestly trying to further science, while only a handful are more motivated by career-maximization. If scholars get promoted based on their most famous papers, the possibility that their best papers would be the focus of a replication attempt should heighten authors' desire to be sure that what they publish is air-tight.

Our primary motive, however, was that we expected that readers would be more interested in learning about the scope and limits of important findings than of unimportant ones. Since we started the replication corner, the single biggest reason for rejecting submissions has been that the original paper did not meet our threshold for importance and influence, as reflected in awards and citations relative to other findings published in top marketing journals in the same year, augmented by our own judgments of importance.

4. Conceptual rather than "direct" replications

Psychologists have debated whether to promote "direct" or "conceptual" replications. If replications are to fulfill an auditing function, one should maximize the similarity between the procedures used in the original study, relying on so called "direct" replications (Pashler & Harris, 2012; Simons, 2014) — sometimes heroically called "exact" replications. Brandt et al. (2014) have gone so far as to put forth a "replication recipe" for "close" replications.

As we explained earlier, we don't believe that "exact" replications are possible. In any given study, researchers make dozens of decisions about small details of procedure, participants, stimuli, setting, and time. When researchers hold those presumed-irrelevant factors constant, it is not possible to say whether any observed treatment effect is a "main effect" of the treatment or a "simple effect" of the treatment in some interaction with one or more of those background factors (Lynch, 1982). Obviously, many things will always differ between the original study and the attempted replication (Stroebe & Strack, 2014). If some replication attempt fails to detect an effect in the original paper, one cannot say whether the issue is one of statistical conclusion validity of the original authors — the original result was a type I error — or an issue of external validity (Open Science Collaboration, 2015; Shadish, Cook, & Campbell, 2002).

Replication efforts in psychology are often motivated by questions of statistical conclusion validity. They seek to investigate whether the original finding was a type I error, or more demandingly, whether the effect size in the replication matches that in the original. As Meehl (1967) argued persuasively, the null hypothesis is never true. Any treatment has some effect however small, so if one has a directional hypothesis and very large sample size, the probability of getting a significant result in the predicted direction approaches 0.5.

Suppose the effect in the population is nonzero but tiny. If the replication doesn't find a statistically significant result, a contributing factor may arise from a publication bias for significant findings. Any such publication bias makes it likely that the effect size in the original (published) study is greater than the effect size in the population from which the original study was drawn. That is exactly what was found in the Open Science Collaboration (2015) report of attempts to replicate 100 psychology articles. We don't find such shrinkage surprising: it is a logical consequence of regression to the mean in a system with publication bias that censors the reporting of small effect sizes. Such shrinkage reflects poorly on our publication system, but not on the individual authors.

Moreover, if replicators base power analyses on inflated effect size estimates from published studies, they may have far less power than their calculations imply. So failure to find a significant result in the replication is less probative of whether an original result was a type I error than some replicators imagine. All of this makes us skeptical of using replications to sort out issues of statistical conclusion validity in the original study.[2]

We, the co-editors of the Replication Corner, believe that it is more interesting to investigate the external validity (robustness) of an influential study. To what extent are the sign, significance, and effect size of original result robust with respect to changes in the stimuli, settings, participant characteristics, contexts, and time of the study? We believe that conceptual replications are more informative than "direct"

[2] Here, our skepticism does not extend to efforts to sort out issues of statistical conclusion validity by placing p values or effect sizes in some distribution of test statistics. For example, Simonsohn, Simmons and Nelson (2015) have developed a "specification curve" methodology for determining what percent of all plausible specifications of an empirical/econometric model would reproduce the sign and the significance of the effect reported in the original paper. If it can be shown that only a very small fraction of plausible specifications produce an effect of the same sign or significance, this would raise the plausibility of questions about statistical conclusion validity.

Please cite this article as: Lynch, J.G., et al., Reflections on the replication corner: In praise of conceptual replications, *IJRM* (2015), http://dx.doi.org/10.1016/j.ijresmar.2015.09.006

replications if the objective is to better understand the external validity or generalizability of the finding. Information on background factor by treatment interactions helps us better understand the nomological validity of the theory that was used to interpret the original finding (Lynch, 1982, 1983).

4.1. Consequences of "successful" direct vs. conceptual replication

Consider the consequences of "successful" replication and "failures to replicate" under two possible replication strategies:

(a) when the authors attempt "direct" replication by matching the replicate to the original study on as many background factors as possible;
(b) when authors attempt "conceptual" replication, replicating the conceptual independent and dependent variables with operationalizations that vary in multiple ways from the original and with multiple background factors held constant at levels different from the original.

The latter strategy corresponds to what Cook and Campbell (1979) have called "deliberate sampling for heterogeneity."

We believe that one learns more from a "successful" conceptual replication than from a successful direct replication. In the case of a direct replication, it may be the case that the results derive from shared interactions with the various background factors that have been held constant in the original and the replication. In the case of a successful "conceptual" replication, it becomes much less plausible that the "main" effect of the treatment in question is confounded with some background factor by treatment interaction (Lynch, 1982, 1983). From a Bayesian perspective, this increases the amount of belief shift warranted by the combined original study and "successful" replication (Brinberg, Lynch, & Sawyer, 1992).

As we will report later in this paper, the replications we published largely reproduced the original authors' effects. We believe that less would have been learned about the broad phenomena studied if the same papers had faithfully attempted to follow exactly the original authors' methods.

4.2. Consequences of "unsuccessful" direct vs. conceptual replication

Now consider the same comparison when the replication attempt fails to reproduce the original effect. The first thing the replicator should do is to see if the inconsistency in results exceeds what might be expected by chance. As noted on the IRJM Website (http://portal.idc.ac.il/en/main/research/ijrm/pages/submission-guidelines.aspx):

"If the original author reports a significant effect and a second author finds no significant effect, it is always unclear whether the difference in results is a "failure to replicate" or just what one would expect from random draws from a common effect size distribution. We would like to ask authors to include in their papers or at least an online appendix a meta-analysis including the original study and attempted replication. An example of such a meta-analysis in the online appendix comes from Chark and Muthukrishnan (IJRM Dec 2013). It is easy to have a situation where one effect size is significant and another is not, but no significant heterogeneity exists across the studies. If the heterogeneity is not significant, then one can calculate the weighted average effect size and test whether the effect is significant after pooling across all the available studies. If there is no significant heterogeneity but the weighted average effect size remains significant, the original conclusion would stand. If there is no significant heterogeneity and the weighted average effect size is NOT significant, then this calls into question the original finding. If there is significant heterogeneity, then this raises the question of what is the moderator or boundary condition that explains the difference in results."

We are surprised that in psychology and economics, it does not seem to be common practice for the authors of an unsuccessful replication to conduct such a meta-analysis, and we are pleased to see that in the final report of the Open Science Collaboration (2015), meta-analytic tests of aggregate effect sizes were reported. In 2014, we began to encourage such analysis in the IJRM Replication Corner, where feasible.

In the case where there is no significant heterogeneity of effect sizes and the combined effect is not significant, direct and conceptual replications are similar. In both cases, a Bayesian process of belief revision will often lead to lower belief in the construct-to-construct links asserted in the original paper. This is a form of learning because before the failed replication, one believed the original effect and its interpretation and now the contrary results put that into question. This is contrary to the narrow definition of learning in which posterior uncertainty is reduced by new data; that is, in this case we "learn we know less" (Bradlow, 1996).

In the case in which the replicator finds no effect and that effect is significantly different from the original in meta-analytic tests, readers of the report wouldn't understand why the results of the original and replicate differed without further data — no matter whether it is a conceptual or a direct replication. It is to be expected that inferences from unsuccessful replications are not definitive but require more research. One might take the position that a "failed" replication is more informative if it is an "exact" replication versus one that differs from the original on multiple dimensions. This is not true if the goal is to establish broad empirical generalizations that may require multiple studies and meta-analysis. This suggests that any given study is (just) one data point for a larger meta-analysis involving future studies. If one is anticipating meta-analysis after more findings accumulate, one should be thinking in terms of what next study might successfully discriminate between plausible alternative causes of variations

Please cite this article as: Lynch, J.G., et al., Reflections on the replication corner: In praise of conceptual replications, *IJRM* (2015), http://dx.doi.org/10.1016/j.ijresmar.2015.09.006

in effect size and most increase our understanding of key moderators and boundary conditions (Farley, Lehmann, & Mann, 1998).

4.3. Failures to replicate are not shameful for the original authors

In psychology, failures to replicate are often taken to shame the original authors, inspiring acrimony about the degree to which the replicators have faithfully followed the original. We think this is unfortunate. Cronbach (1975) has argued persuasively that most real world behavior is driven by higher-order interactions that are virtually impossible to anticipate. He gives the following example:

> "Investigators checking on how animals metabolize drugs found that results differed mysteriously from laboratory to laboratory. The most startling inconsistency of all occurred after a refurbishing of a National Institutes of Health (NIH) animal room brought in new cages and new supplies. Previously, a mouse would sleep for about 35 min after a standard injection of hexobarbital. In their new homes, the NIH mice came miraculously back to their feet just 16 min after receiving a shot of the drug. Detective work proved that red-cedar bedding made the difference, stepping up the activity of several enzymes that metabolize hexobarbital. Pine shavings had the same effect. When the softwood was replaced with birch or maple bedding like that originally used, drug response came back in line with previous experience" (p. 121).

Who would ever be smart enough to anticipate that one? Our view is that if a colleague's finding is not replicated in an attempted "direct" replication, it is not (necessarily or even usually) a sign of something underhanded or sloppy. It simply means that in the current state of knowledge, we may not fully understand the effect or what moderates it (Cesario, 2014; Lynch, 1982).

4.4. Hidden background factors influence effect sizes

The journal *Social Psychology* published a report by the Many Labs Project (Klein et al., 2014), wherein teams of researchers at 36 different universities attempted direct replications of 16 studies from 13 important papers. In aggregate they successfully replicated 10 of the 13 papers. The teams of researchers at the 36 different universities all followed the same pre-registered protocol. Nonetheless, meta-analytic Q and I^2 statistics showed substantial and statistically significant unexplained heterogeneity across labs for 8 of the 16 effects studied.

In theoretical research, the objective is often to make claims about construct-to-construct links. When one researcher fails to replicate some original finding, it is possible that the replicate and original don't differ in construct-to-construct links; rather, the original and the replicate may differ in the mapping from operational variables to latent constructs. One route to this outcome is when the replicate and original differ in characteristics of participants (e.g. Aaker & Maheswaran, 1997).

The mapping from operational to latent variables can also change with time. A direct replication of a 30 year old study might be technically equivalent in terms of operational independent and dependent variables but with differences in the conceptual meaning of the same operational variables (e.g. Schwarz & Strack, 2014, p. 305). Just as construct-to-construct relations can change over time (Gergen, 1973; Schooler, 2011), so too can operationalization-to-construct links.

In psychology, reports of failures to replicate important papers have been prosecutorial, as if the original effect was not "real." For example, Johnson, Cheung, and Donnellan (2014) failed to replicate a finding by Schnall, Benton, and Harvey (2008) that washing hands led to more lenient moral judgments. Donnellan's (2014) related blog post implied that the original effect does not exist.

> "So more research is probably needed to better understand this effect [Don't you just love Mom and Apple Pie statements!]. However, others can dedicate their time and resources to this effect. We gave it our best shot and pretty much encountered an epic fail as my 10-year-old would say. My free piece of advice is that others should use very large samples and plan for small effect sizes."

Donnellan later apologized for the "epic fail" line, but it reflects an underlying attitude shared by other replicators of — "our findings are correct and the original authors' are wrong." That's hubris: both findings are relevant data points in an as-yet-unsolved puzzle. The same issue of *Social Psychology* that published Johnson, Cheung, and Donnellan (2014) reported two separate direct replications of Shih, Pittinsky, and Ambady's (1999) famous finding that Asian-American women performed better on a mathematics test when their ethnic identity was activated, but worse when their gender identity was activated. Gibson, Losee, and Vitiello (2014) replicated the original finding, but Moon and Roeder (2014) did not — despite following the same pre-registered protocol. If one should be embarrassed to have another lab fail to replicate one's own original result, how should we feel when two different "direct" replications produce different results? And how should the researchers in the "Many Labs Replication Project" (Klein et al., 2014) feel about the fact that their colleagues at other universities found different effect sizes when following the same pre-registered replication research protocols?

In summary, successful direct replications may strengthen our prior beliefs about a well-known effect, but not as much as successful conceptual replications. When replicators find results different from the original, direct replications are like conceptual replications in requiring future research to understand the reasons for differences. Worse, a focus on direct replications can create an unhealthy atmosphere in our field where the competence or honesty of researchers is subtly questioned. Variation in results is to be expected. The next section reveals that it has been a challenge for us to deal with defensiveness on the part of authors whose findings did not replicate.

Please cite this article as: Lynch, J.G., et al., Reflections on the replication corner: In praise of conceptual replications, *IJRM* (2015), http://dx.doi.org/10.1016/j.ijresmar.2015.09.006

5. Even-handed publication of successful conceptual replications and of failures to replicate

We have proposed that the original authors should not feel embarrassed when another author team fails to replicate their results. But in our experience, that is not how original authors often see it. As editors, when we have received a failure-to-replicate report, we have commonly included one of the original authors on the review team for that paper. It is not uncommon for the reviewing author to be a bit defensive, pointing out differences of the replication and the original as flaws.

We have tried hard as editors to push back against such defensiveness, though we are not sure we have been completely successful. As we will lay out in the next section, the percentage of successful replications that we have published seems comparable to what was reported in the Many Labs test of 16 effects from 13 famous papers (Klein et al., 2014). However, the percent of successful conceptual replications in the Replication Corner significantly exceeds the percentage of successful direct replications in the Reproducibility Project (Open Science Collaboration, 2015). The Reproducibility Project is the largest and most ambitious open source replication effort to date, involving 250 scientists around the world. Collectively, they attempted to replicate 100 studies published in the 2008 volumes of *Psychological Science*, *Journal of Personality and Social Psychology*, and *Journal of Experimental Psychology: Learning, Memory and Cognition*. A summary of that effort classified 61 findings as not replicated to varying degrees and only 39 were replicated to varying degrees (Baker, 2015). That's a far lower rate for successful direct replications than we are observing for conceptual replications.

The lesson we derive as editors of the Replication Corner is that we need to be even more aggressive in pushing back when original authors recommend rejection of unsuccessful replications of their work. Pashler and Harris (2012) have correctly noted that absent any systematic replication effort, the normal journal review process makes it more likely that successful conceptual replications will be published than unsuccessful ones because successful conceptual replications seem more interesting. A journal section dedicated to replication should avoid that bias.

6. What we have found so far

Thus far we have published or accepted 30 replications from 91 submissions. Table 1 summarizes the nature of the differences between the replicated articles and the conceptual replications and extensions in those 30 papers. We evaluated these articles on four dimensions shown in Table 2:

1. Direct replication included? Did the authors include at least a subset of experiment design cells/subjects intended to have a very similar operationalization to the original: 1 = yes; 0 = no, conceptual replication only.
2. Moderator? Did the paper show an interaction in which the original result replicates in some conditions but contains different effects under other conditions? (1 = yes; 0 = no)
3. Resolve conflict? Does the paper address and resolve apparent conflicts in the literature? (1 = yes; 0 = no)
4. How closely did the findings agree with the original study or studies? (From Baker, 2015), 1 = Not at all similar; 2 = slightly similar; 3 = somewhat similar; 4 = moderately similar; 5 = very similar; 6 = extremely similar; 7 = virtually identical.

Table 2 shows that most papers reported replicating the original findings to a large degree. The mean of two coders on our 7 point scale was 5.8 on a scale where 6 means "extremely similar" (Cronbach's $\alpha = .76$). Only three papers were coded as less than "4 = very similar" to the original: Aspara & van den Berg (2014), replicating Alexander (2006); Baxter, Kulczynski, & Ilicic (2014), replicating Yorkston & Menon (2004); Gill & El Gamal (2014) replicating Berger & Fitzsimons (2008).

Of the 30 articles, 22 were exclusively conceptual replications; only 8 included at least some conditions intended to somewhat closely match operationalizations of the original authors. Twelve of the studies showed some moderation of the original findings. Seven were coded as showing a resolution of some inconsistency in the literature. As an example Müller, Schliwa, and Lehmann (2014) replicated Simonson and Tversky's (1992) prize decoy experiment that Frederick, Lee, and Baskin (2014) had been unable to replicate. They followed Simonson's (2014) reply to Frederick et al. using real and not hypothetical gambles with asymmetrically dominated (rather than truly inferior) decoys. This echoes our earlier point that failures to replicate can themselves fail to replicate.

Because this coding was not done with the rigor of a formal content analysis, we emailed a draft of the paper with our tentative codes to the authors of the 30 replications; all 30 replied. We received very few minor corrections, and in Table 2, we deferred to the authors in those cases.

7. Conclusion

We are proud to have served as co-editors of the Replication Corner under IJRM editors Jacob Goldenberg and Eitan Muller. We believe that the findings published in the Replication Corner have had a distinctly positive influence on the field of marketing, serving to enhance the sense that our most important findings are not perilously fragile.

The incoming editor of IJRM will discontinue the Replication Corner in the face of the reality of journal rankings based on citation impact factors. Replication papers are crucial for the field, but on average they may be cited less than the regular-length articles in the same journals. We were heartened to learn that the new EMAC *Journal of Marketing Behavior* has eagerly agreed to continue the Replication Corner. *JMB* editor Klaus Wertenbroch took this decision although the Reproducibility Project did not replicate findings from Dai, Wertenbroch, and Brendl (2008). We infer that Wertenbroch

Please cite this article as: Lynch, J.G., et al., Reflections on the replication corner: In praise of conceptual replications, *IJRM* (2015), http://dx.doi.org/10.1016/j.ijresmar.2015.09.006

Table 1
Summary of nature of replications and extensions of 30 replication corner papers.

Authors	Title	Paper Replicated	Nature of extension
Aspara and van den Berg (2014)	Naturally designed for masculinity vs. femininity? Prenatal testosterone predicts male consumers' choices of gender-imaged products	Alexander (2006), Archives of Sexual Behavior	Tested consequences of digit ratios for new DV: Preference for gender-linked products.
Barakat et al. (2015)	Severe service failure recovery revisited: Evidence of its determinants in an emerging market context	Weun et al. (2004), J. Services Marketing	Extended prior findings on link of service failure severity to satisfaction to new emerging market and new industry. Extended replicated paper by testing how three perceived justice dimensions moderate that relationship in real (rather than experimental) service encounters.
Baxter and Lowrey (2014)	Examining children's preference for phonetically manipulated brand names across two English accent groups	Shrum et al., 2012 International J. of Research in Marketing	Tested moderation of brand sound preferences by accent and age
Baxter et al. (2014)	Revisiting the automaticity of phonetic symbolism effects	Yorkston & Menon, 2004. J. Consumer Research	Tested moderation of automatic phonetic symbolism effects across adults and children
Blut et al. (2015)	How procedural, financial and relational switching costs affect customer satisfaction, repurchase intentions, and repurchase behavior: A meta-analysis	Burnham et al. (2003), J. of the Academy of Marketing Science	Meta-analysis of conflicting studies on effects of satisfaction and switching costs on repurchase behavior. Examined moderation by DV of intentions vs. actual behavior.
Brock et al. (2013)	Satisfaction with complaint handling: A replication study on its determinants in a business-to-business context	Orsingher et al. (2010), J. of the Academy of Marketing Science.	Extend original findings into B to B context
Butori and Parguel (2014)	The impact of visual exposure to a physically attractive other on self-presentation	Roney (2003), Personality and Social Psychology Bulletin	Extended by use of different stimuli (pictures in a non-mating context, whereas previous studies used pictures in a mating context), and analyzed a different population : women (and not just men).
Chan (in press-a)	Attractiveness of options moderates the effect of choice overload	Gourville and Soman (2005), Marketing Science; Iyengar & Lepper (2000, J. Personality & Social Psychology)	Examined a moderating variable for the choice overload effect, namely how attractive the options in a choice set are
Chan (in press-b)	Endowment effect for hedonic but not utilitarian goods	Ariely et al. (2005), J. Marketing Research; Kahneman et al. (1990) J. Political Economy.	Examined a moderating variable for the endowment effect, namely by comparing hedonic against utilitarian goods
Chark and Muthukrishnan (2013)	The effect of physical possession on preference for product warranty	Peck and Shu (2009), J. Consumer Research	Generarlized effect of physical contact from perceived ownership to intention to buy extended warranties
Chowdhry et al. (2015)	Not all negative emotions lead to concrete construal	Labroo and Patrick (2009), J. Consumer Research	Extended the original research by showing that appraisals of specific emotions (rather than valence alone) impacts construal level
Davvetas, Sichtman & Diamantopoulos (in press)	The impact of perceived brand globalness on consumers' willingness to pay	Steenkamp et al. (2003), J. International Business Studies	Manipulated rather than measured brand globalness, tested multiple moderators and found few significant, replicating in new product categories
Evanschtzki et al. (2014)	Hedonic shopping motivations in collectivistic and individualistic consumer cultures	Arnold and Reynolds (2003), J. Retailing	Replicated original results in individualistic cultures, but showed different effects of shopping motivations in collectivist cultures
Fernandes (2013)	The 1/N rule revisited: Heterogeneity in the naïve diversification bias	Benartzi and Thaler (2001), American Econ. Review	Tested and refuted two previous explanations for diversification bias, desire for variety and financial knowledge, and showed role of reliance on intuition in diversification bias
Gill and El Gamal (2014)	Does exposure to dogs (cows) increase the preference for puma (the color white)? Not always	Berger and Fitzsimons (2008), J. Marketing Research	Extended test for priming effects to different stimuli, different population sample (general population, plus students in the lab), and different frequencies of exposure
Hasford, Farmer, & Waites (in press)	Thinking, feeling, and giving: The effects of scope and valuation on consumer donations	Hsee and Rottenstreich (2004), J. Experimental Psychology: General	Used new charity for tests and showed moderation by understanding of emotional intelligence, shedding light on mechanism for scope insensitivity
Holden and Zlatevska (2015)	The partitioning paradox: The big bite around small packages	Do Vale et al. (2008), J. Consumer Research; Scott et al. (2008), J. Consumer Research	Blended methods from three prior studies, testing in different country and shorter time period. Extended by showing that respondent awareness of participation in a food study eliminated the effect.
Holmqvist and Lunardo (2015)		Kaltcheva and Weitz (2006), J. Marketing	Extended to new culture and tested mediators of original effect

(continued on next page)

Please cite this article as: Lynch, J.G., et al., Reflections on the replication corner: In praise of conceptual replications, *IJRM* (2015), http://dx.doi.org/10.1016/j.ijresmar.2015.09.006

Table 1 (*continued*)

Authors	Title	Paper Replicated	Nature of extension
	The impact of an exciting store environment on consumer pleasure and behavioral intentions		
Huyghe and van Kerckhove (2013)	Can fat taxes and package size restrictions stimulate healthy food choices?	Mishra and Mishra (2011), J. Marketing Research	Extended to choices between healthy and unhealthy foods. Showed that price changes for indulgent option but not healthy option affected intentions to buy the indulgent option. Changes in package size of the healthy but not the indulgent option affected intentions to purchase the indulgent option.
Kuehnl and Mantau (2013)	Same sound, same preference? Investigating sound symbolism effects in international brand names	Lowrey & Shrum (2007) J. Consumer Research; Shrum et al. (2012), IJRM	Compared sound perceptions and preferences in French, German, and Spanish stimuli. Created new fictitious brand names and added consideration of the effect of consonants in brand names
Lenoir et al. (2013)	The impact of cultural symbols and spokesperson identity on attitudes and intentions	Deshpandé and Stayman (1994), J. Marketing Research; Forehand and Deshpandé (2001), J. Marketing Research	Showed targeted marketing strategies that work for first generation minority consumers do not work for second generation minority consumers & vice versa
Lin (2013)	Does container weight influence judgments of volume?	Krishna (2006), J. Consumer Research	Replicated finding that longer cylinders are perceived to have more volume than shorter ones of equal volume, then showed this bias goes away when weight cues are incorporated into volume judgments
Maecker et al. (2013)	Charts and demand: Empirical generalizations on social influence	Salganik et al. (2006), Science	Different subject times, stimuli, and product classes
Mukherjee (2014)	How chilling are network externalities? The role of network structure	Goldenberg et al. (2010), IJRM	Across 7 real world data sets, author demonstrated that the conclusion that externalities slow adoption is not a tautological consequence of original model formulation; that higher size and higher average degree can offset the effect of network externalities, and that more clustering in the network strengthens the chilling effect of externalities
Müller (2013)	The real-exposure effect revisited - How purchase rates vary under pictorial vs. real item presentations when consumers are allowed to use their tactile sense	Bushong et al. (2010), Am. Econ. Rev.	Extended to different modes of real exposure; purchase DV vs. Becker Degroot Marshak; appetitive vs. nonappetitive goods; high vs. low familiarity
Müller et al. (2014)	Prize decoys at work — New experimental evidence for asymmetric dominance effects in choices on prizes in competitions	Simonson and Tversky (1992), J. Marketing Research; Frederick et al. (2014), J. Marketing Research	Extended to real consequential choices among options tested to produce tradeoffs claimed to be necessary for asymmetric dominance effect
Müller et al. (2013)	The time vs. money effect. A conceptual replication	Mogilner and Aaker (2009), J. Consumer Research	Extended from field to lab, tested treatment interactions with demographic background variables
Orazi and Pizzetti (2015)	Revisiting fear appeals: A structural re-Inquiry of the protection motivation model	Johnston and Warkentin (2010), MIS Quarterly	Conceptuallly replicated original with different subject types (adults), product types (online banking security), and model specification and estimation
Van Doorne et al. (2013)	Satisfaction as a predictor of future performance: A replication	Keiningham et al. (2007), JM; Morgan and Rego (2006), Marketing Science; Reichheld (2003)HBR	Prior work disputed which customer metric best predicts future company performance. Authors assessed the impact of different satisfaction and loyalty metrics as well as the Net Promoter Score on sales revenue growth, gross margins and net operating cash flows using a Dutch sample.
Wright et al. (2013)	If it tastes bad it must be good: Consumer naïve theories and the marketing placebo effect	Shiv et al. (2005), J. Marketing Research	Replicated original price placebo effect using unique stimuli, subject types, and dependent variables. Showed that effect extends to other cues: Set size, product typicality, product taste, and shelf availability.

shares our view that "unsuccessful" and "successful" replications are a valuable contribution to science and that there is no personal affront when another scholar reports an "unsuccessful" replication of one's earlier findings.

Like Klaus Wertenbroch, we are committed to the cause and will follow the Replication Corner to its new home at *Journal of Marketing Behavior*. We hope that readers of this editorial will similarly continue to support the Replication Corner.

Please cite this article as: Lynch, J.G., et al., Reflections on the replication corner: In praise of conceptual replications, *IJRM* (2015), http://dx.doi.org/10.1016/j.ijresmar.2015.09.006

Table 2
Coding of papers appearing in replication corner.

Authors	Title	Paper replicated	Direct replication included? (1 = Yes; 0 = No)	Moderation of effect shown? (0 = No; 1 = Yes)	Resolve conflict among papers? (1 = Yes; 0 = No)	Replication score (1 = Not At all similar; 7 = virtually identical)
Aspara and van den Berg (2014)	Naturally designed for masculinity vs. femininity? Prenatal testosterone predicts male consumers' choices of gender-imaged products	Alexander (2006), Archives of sexual behavior	0	0	0	3
Barakat, Ramsey, Lorenz, and Gosling (2015)	Severe service failure recovery revisited: Evidence of its determinants in an emerging market context	Weun, Beatty, and Jones (2004), J. Services Marketing	0	0	1	5.5
Baxter and Lowrey (2014)	Examining children's preference for phonetically manipulated brand names across two english accent groups	Shrum, Lowrey, Luna, Lerman, & Liu, 2012 International J. of Research in Marketing	0	1	0	5
Baxter et al. (2014)	Revisiting the automaticity of phonetic symbolism effects	Yorkston & Menon, 2004. J. Consumer Research	1	0	0	2.5
Blut, Frennea, Mittal, and Mothersbaugh (2015)	How procedural, financial and relational switching costs affect customer satisfaction, repurchase intentions, and repurchase behavior: A meta-analysis	Burnham, Frels, and Mahajan (2003), J. of the Academy of Marketing Science	0	1	1	5.5
Brock, Blut, Evanschitzky, and Kenning (2013)	Satisfaction with complaint handling: A replication study on its determinants in a business-to-business context	Orsingher, Valentini, and deAngelis (2010), J. of the Academy of Marketing Science.	0	0	0	5
Butori and Parguel (2014)	The impact of visual exposure to a physically attractive other on self-presentation	Roney (2003), Personality and Social Psychology Bulletin	0	0	0	7
Chan (in press-a)	Attractiveness of options moderates the effect of choice overload	Gourville and Soman (2005), Marketing Science; Iyengar & Lepper (2000, J. Personality & Social Psychology)	0	1	1	7
Chan (in press-b)	Endowment effect for hedonic but not utilitarian goods	Ariely, Huber, and Wertenbroch (2005), J. Marketing Research; Kahneman, Knetsch, and Thaler (1990), J. Political Economy.	0	1	0	7
Chark and Muthukrishnan (2013)	The effect of physical possession on preference for product warranty	Peck and Shu (2009), J. Consumer Research	0	0	0	7
Chowdhry, Winterich, Mittal, and Morales (2015)	Not all negative emotions lead to concrete construal	Labroo and Patrick (2009), J. Consumer Research	1	0	0	6.5
Davvetas, Sichtmann, & Diamantopoulos (in press)	The impact of perceived brand globalness on consumers' willingness to pay	Steenkamp, Batra, and Alden (2003), J. International Business Studies	0	0	0	5
Evanschtzki et al. (2014)	Hedonic shopping motivations in collectivistic and individualistic consumer cultures	Arnold and Reynolds (2003), J. Retailing	1	1	0	7
Fernandes (2013)	The 1/N rule revisited: Heterogeneity in the naïve diversification bias	Benartzi and Thaler (2001), American Econ. Review	1	1	1	6
Gill and El Gamal (2014)	Does exposure to dogs (cows) increase the preference for puma (the color white)? Not always	Berger and Fitzsimons (2008), J. Marketing Research	0	0	0	3.5
			0	1	0	7

(continued on next page)

Table 2 (*continued*)

Authors	Title	Paper replicated	Direct replication included? (1 = Yes; 0 = No)	Moderation of effect shown? (0 = No; 1 = Yes)	Resolve conflict among papers? (1 = Yes; 0 = No)	Replication score (1 = Not At all similar; 7 = virtually identical)
Hasford, Farmer, & Waites (in press)	Thinking, feeling, and giving: The effects of scope and valuation on consumer donations	Hsee and Rottenstreich (2004), J. Experimental Psychology: General				
Holden and Zlatevska (2015)	The partitioning paradox: The big bite around small packages	Do Vale, Pieters, and Zeelenberg (2008), J. Consumer Research; Scott, Nowlis, Mandel, and Morales (2008), J. Consumer Research	0	1	0	6.5
Holmqvist and Lunardo (2015)	The impact of an exciting store environment on consumer pleasure and behavioral intentions	Kaltcheva and Weitz (2006), J. Marketing	1	0	0	5
Huyghe and van Kerckhove (2013)	Can fat taxes and package size restrictions stimulate healthy food choices?	Deshpandé and Stayman (1994), J. Marketing Research; Forehand and Deshpandé (2001), J. Marketing Research	0	0	0	6.5
Kuehnl and Mantau (2013)	Same sound, same preference? Investigating sound symbolism effects in international brand names	Lowrey & Shrum (2007) J. Consumer Research; Shrum et al. (2012), IJRM	0	0	0	5.5
Lenoir, Puntoni, Reed, and Verlegh (2013)	The impact of cultural symbols and spokesperson identity on attitudes and intentions	Deshpandé and Stayman (1994), J. Marketing Research; Forehand and Deshpandé (2001), J. Marketing Research	0	1	0	5
Lin (2013)	Does container weight influence judgments of volume?	Krishna (2006), J. Consumer Research	0	1	0	6
Maecker, Grabenströer, Clement, and Heitmann (2013)	Charts and demand: Empirical generalizations on social influence	Salganik, Dodds, and Watts (2006), Science	1	0	0	6.5
Mukherjee (2014)	How chilling are network externalities? The role of network structure	Goldenberg, Libai, and Muller (2010), IJRM.	1	1	1	6
Müller (2013)	The real-exposure effect revisited - How purchase rates vary under pictorial vs. real item presentations when consumers are allowed to use their tactile sense	Bushong, King, Camerer, and Rangel (2010), Am. Econ. Rev.	0	1	0	7
Müller et al. (2014)	Prize decoys at work — New experimental evidence for asymmetric dominance effects in choices on prizes in competitions	Simonson and Tversky (1992), J. Marketing Research; Frederick et al. (2014), J. Marketing Research	0	0	1	6.5
Müller, Lehmann, and Sarstedt (2013)	The time vs. money effect. A conceptual replication	Mogilner and Aaker (2009), J. Consumer Research	1	0	0	7
Orazi and Pizzetti (2015)	Revisiting fear appeals: A structural re-Inquiry of the protection motivation model	Johnston and Warkentin (2010), MIS Quarterly	0	0	0	5.5
Van Doorne, Leeflang, and Tijs (2013)	Satisfaction as a predictor of future performance: A replication	Keiningham, Cooil, Andreasson, and Aksoy (2007), JM; Morgan and Rego (2006), Marketing Science; Reichheld (2003), HBR	0	0	1	5.5
Wright, Hernandez, Sundar, Dinsmore, and Kardes (2013)	If it tastes bad it must be good: Consumer naïve theories and the marketing placebo effect	Shiv, Carmon, and Ariely (2005), J. Marketing Research	0	0	0	6.5

Please cite this article as: Lynch, J.G., et al., Reflections on the replication corner: In praise of conceptual replications, *IJRM* (2015), http://dx.doi.org/10.1016/j.ijresmar.2015.09.006

References

Aaker, J. L., & Maheswaran, D. (1997). The effect of cultural orientation on persuasion. *Journal of Consumer Research*, *24*(3), 315–328.

Alexander, G. M. (2006). Associations among gender-linked toy preferences, spatial ability,and digit ratio: Evidence from eye-tracking analysis. *Archives of Sexual Behavior*, *35*(6), 699–709.

Ariely, D., Huber, J., & Wertenbroch, K. (2005). When do losses loom larger than gains? *Journal of Marketing Research*, *42*, 134–138.

Ariely, D., Loewenstein, G., & Prelec, D. (2003). Coherent arbitrariness: Stable demand curves without stable preferences. *The Quarterly Journal of Economics*, *118*(1), 73–105.

Arnold, M. J., & Reynolds, K. E. (2003). Hedonic shopping motivations. *Journal of Retailing*, *79*(2), 77–95. , http://dx.doi.org/10.1016/S0022-4359(03)00007-1.

Aspara, J., & van den Berg, B. (2014). Naturally designed for masculinity vs. femininity? Prenatal testosterone predicts male consumers' choices of gender-imaged products. *International Journal of Research in Marketing*, *31*(1), 117–121.

Baker, M. (2015, April 30). First results from psychology's largest reproducibility test: Crowd-sourced effort raises nuanced questions about what counts as a replication. *Nature*, 2015. , http://dx.doi.org/10.1038/nature.2015.17433.

Barakat, L. L., Ramsey, J. R., Lorenz, M. P., & Gosling, M. (2015). Severe service failure recovery revisited: Evidence of its determinants in an emerging market context. *International Journal of Research in Marketing*, *32*(1), 113–116.

Bargh, J., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology*, *71*(2), 230–244.

Baxter, S., & Lowrey, T. M. (2014). Examining children's preference for phonetically manipulated brand names across two English accent groups. *International Journal of Research in Marketing*, *31*(1), 122–124.

Baxter, S., Kulczynski, A., & Ilicic, J. (2014). Revisiting the automaticity of phonetic symbolism effects. *International Journal of Research in Marketing*, *31*(4), 448–451.

Benartzi, S., & Thaler, R. H. (2001). Naive diversification strategies in retirement saving plans. *American Economic Review*, *91*(1), 79–98.

Berger, J., & Fitzsimons, G. (2008, February). Dogs on the street, puma on the feet: How cues in the environment influence product evaluation and choice. *Journal of Marketing Research*, *45*, 1–14.

Blut, M., Frennea, C. M., Mittal, V., & Mothersbaugh, D. L. (2015). How procedural, financial and relational switching costs affect customer satisfaction, repurchase intentions, and repurchase behavior: A meta-analysis. *International Journal of Research in Marketing*, *32*(2), 226–229.

Bradlow, E. T. (1996). Negative information and the three-parameter logistic model. *Educational and Behavioral Statistics*, *21*(2), 179–185.

Brandt, M. J., IJzerman, H., Dijksterhuis, A., Farach, F. J., Geller, J., Giner-Sorolla, R., ... Van't Veer, A. (2014,). The replication recipe: What makes for a convincing replication? *Journal of Experimental Social Psychology*, *50*, 217–224.

Brinberg, D. L., Lynch, J. G., Jr., & Sawyer, A. G. (1992, September). Hypothesized and confounded explanations in theory tests: A Bayesian analysis. *Journal of Consumer Research*, *19*, 139–154.

Brock, C., Blut, M., Evanschitzky, H., & Kenning, P. (2013). Satisfaction with complaint handling: A replication study on its determinants in a business-to-business context. *International Journal of Research in Marketing*, *30*(3), 319–322.

Burnham, T. A., Frels, J. K., & Mahajan, V. (2003). Consumer switching costs: A typology, antecedents, and consequences. *Journal of the Academy of Marketing Science*, *31*(2), 109–127.

Bushong, B., King, L. M., Camerer, C. F., & Rangel, A. (2010). Pavlovian processes in consumer choice: The physical presence of a good increases willingness-to-pay. *American Economic Review*, *100*(4), 1556–1571.

Butori, R., & Parguel, B. (2014). The impact of visual exposure to a physically attractive other on self-presentation. *International Journal of Research in Marketing*, *31*(3), 445–447.

Cesario, J. (2014). Priming, replication, and the hardest science. *Perspectives on Psychological Science*, *9*(1), 40–48.

Chan, E. Y. (2015a). Attractiveness of options moderates the effect of choice overload. *International Journal of Research in Marketing*, *32*(4) (in press).

Chan, E. Y. (2015b). Endowment effect for hedonic but not utilitarian goods. *International Journal of Research in Marketing*, *32*(4) (in press).

Chark, R., & Muthukrishnan, A. V. (2013). The effect of physical possession on preference for product warranty. *International Journal of Research in Marketing*, *30*(4), 424–425.

Chowdhry, N., Winterich, K. P., Mittal, V., & Morales, A. C. (2015). Not all negative emotions lead to concrete construal. *International Journal of Research in Marketing*, *32*(4), http://dx.doi.org/10.1016/j.ijresmar.2015.03.006 (in press).

Coan, J. (2014). Negative psychology: The atmosphere of wary and suspicious disbelief. Blog post in Medium.com https://medium.com/@jimcoan/negative-psychology-f66795952859

Cook, T. K., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings.* Chicago: Rand McNally.

Cronbach, L. J. (1975). Beyond the two disciplines of scientific psychology. *American Psychologist*, *30*, 116–127.

Dai, X., Wertenbroch, K., & Brendl, C. M. (2008). The value heuristic in judgments of relative frequency. *Psychological Science*, *19*(1), 18–19.

Davvetas, V., Sichtman, C., & Diamantopoulos, A. (2015). The impact of perceived brand globalness on consumers' willingness to pay. *International Journal of Research in Marketing*, *32*(4), http://dx.doi.org/10.1016/j.ijresmar.2015.05.004 (in press).

Deshpandé, R., & Stayman, D. M. (1994). A tale of two cities: Distinctiveness theory and advertising effectiveness. *Journal of Marketing Research*, *31*(1), 57–64.

Do Vale, R. C., Pieters, R., & Zeelenberg, M. (2008, October). Flying under the radar: Perverse package size effects on consumption self-regulation. *Journal of Consumer Research*, *35*, 380–390.

Donnellan, D. (2014). Go big or go home: A recent replication attempt. Blog post https://traitstate.wordpress.com/2013/12/11/go-big-or-go-home-a-recent-replication-attempt/

Doyen, S., Klein, O., Pichon, C., & Cleeremans, A. (2012). Behavioral priming: It's all in the mind, but whose mind. *PloS One*, *7*(1), e29081.

Evanschtzki, H., Emrich, O., Sangtani, V., Ackfeldt, L., Reynolds, K. E., & Arnold, M. J. (2014). Hedonic shopping motivations in collectivistic and individualistic consumer cultures. *International Journal of Research in Marketing*, *31*(3), 335–338.

Fang, F. C., Steen, R. G., & Casadevall, A. (2012). Misconduct accounts for the majority of retracted scientific publications. *Proceedings of the National Academy of Sciences*, *109*(42), 17028–17033.

Farley, J. U., Lehmann, D. R., & Mann, L. H. (1998, November). Designing the next study for maximum impact. *Journal of Marketing Research*, *35*, 496–501.

Fernandes, D. (2013). The 1/N rule revisited: Heterogeneity in the naïve diversification bias. *International Journal of Research in Marketing*, *30*(3), 310–313.

Forehand, M. R., & Deshpandé, R. (2001). What we see makes us who we are: Priming ethnic self-awareness and advertising response. *Journal of Marketing Research*, *38*(3), 338–346.

Frederick, S., Lee, L., & Baskin, E. (2014). The limits of attraction. *Journal of Marketing Research*, *51*(4), 487–507.

Gelman, A., & Loken, E. (2014). The statistical crisis in science. *American Scientist*, *102*(6), 460–465.

Gergen, K. J. (1973). Social psychology as history. *Journal of Personality and Social Psychology*, *26*(2), 309.

Gibson, C. E., Losee, J., & Vitiello, C. (2014). A replication attempt of stereotype susceptibility (Shih, Pittinsky, & Ambady, 1999): Identity salience and shifts in quantitative performance. *Social Psychology*, *45*(3), 194–198.

Gilbert, D. (2014). Psychology's replication police prove to be shameless little bullies. Twitter post https://twitter.com/dantgilbert/status/470199929626193921

Gill, T., & El Gamal, M. (2014). Does exposure to dogs (cows) increase the preference for puma (the color white)? Not always. *International Journal of Research in Marketing*, *31*(1), 125–126.

Please cite this article as: Lynch, J.G., et al., Reflections on the replication corner: In praise of conceptual replications, *IJRM* (2015), http://dx.doi.org/10.1016/j.ijresmar.2015.09.006

Goldenberg, J., Libai, B., & Muller, E. (2010). The chilling effects of network externalities. *International Journal of Research in Marketing*, *27*(1), 4–15.

Gourville, J., & Soman, D. (2005). Overchoice and assortment type: When variety backfires. *Marketing Science*, *24*, 382–395.

Hasford, J., Farmer, A., & Waites, S. F. (2015). Thinking, feeling, and giving: The effects of scope and valuation on consumer donations. *International Journal of Research in Marketing*, *32*(4) (in press).

Holden, S. S., & Zlatevska, N. (2015). The partitioning paradox: The big bite around small packages. *International Journal of Research in Marketing*, *32*(2), 230–233.

Holmqvist, J., & Lunardo, R. (2015). The impact of an exciting store environment on consumer pleasure and shopping intentions. *International Journal of Research in Marketing*, *32*(1), 117–119.

Hsee, C. K., & Rottenstreich, Y. (2004). Music, pandas, and muggers: On the affective psychology of value. *Journal of Experimental Psychology: General*, *133*(1), 23. , http://dx.doi.org/10.1037/0096-3445.133.1.23.

Huyghe, E., & van Kerckhove, A. (2013). Can fat taxes and package size restrictions stimulate healthy food choices? *International Journal of Research in Marketing*, *30*(4), 421–423.

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*, *2*(8), e124.

Iyengar, S. S., & Lepper, M. R. (2000). When choice is demotivating: Can one desire too much of a good thing? *Journal of Personality and Social Psychology*, *79*, 995–1006.

Johnson, D. J., Cheung, F., & Donnellan, M. B. (2014). Does cleanliness influence moral judgments? A direct replication of Schnall, Benton, and Harvey (2008). *Social Psychology*, *45*(3), 209–215.

Johnston, A. C., & Warkentin, M. (2010). Fear appeals and information security behaviors: An empirical study. *MIS Quarterly*, *34*(3), 549–566.

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1990). Experimental tests of the endowment effect and the Coase theorem. *Journal of Political Economy*, *98*, 1325–1348.

Kaltcheva, V. D., & Weitz, B. A. (2006). When should a retailer create an exciting store environment? *Journal of Marketing*, *70*(1), 107–118.

Keiningham, T. L., Cooil, B., Andreasson, T. W., & Aksoy, L. (2007). A longitudinal examination of 'net promoter' and firm revenue growth. *Journal of Marketing*, *71*(3), 39–51.

Klein, R. A., Ratliff, K. A., Vianello, M., Adams, R. B., Bahník, S., Bernstein, M. J., Bocian, K., et al. (2014t). Investigating variation in replicability. *Social Psychology*, *45*(3), 142–152.

Krishna, A. (2006). Interaction of senses: the effect of vision versus touch on the elongation bias. *Journal of Consumer Research*, *32*, 557–566.

Kuehnl, C., & Mantau, A. (2013). Same sound, same preference? Investigating sound symbolism effects in international brand names. *International Journal of Research in Marketing*, *30*(4), 417–420.

Labroo, A. A., & Patrick, V. M. (2009). Psychological distancing: Why happiness helps you see the big picture. *Journal of Consumer Research*, *35*(5), 800–809.

Lenoir, A. I., Puntoni, S., Reed, A., & Verlegh, P. W. J. (2013). The impact of cultural symbols and spokesperson identity on attitudes and intentions. *International Journal of Research in Marketing*, *30*(4), 426–428.

Lin, H. -M. (2013). Does container weight influence judgments of volume? *International Journal of Research in Marketing*, *30*(3), 308–309.

Lowrey, T. M., & Shrum, L. J. (2007). Phonetic symbolism and brand name preference. *Journal of Consumer Research*, *34*, 406–414 (October).

Lynch, J. G., Jr. (1982, December). On the external validity of experiments in consumer research. *Journal of Consumer Research*, *9*, 225–239.

Lynch, J. G., Jr. (1983, June). The role of external validity in theoretical research. *Journal of Consumer Research*, *10*, 109–111.

Maecker, O., Grabenströer, N. S., Clement, M., & Heitmann, M. (2013). Charts and demand: Empirical generalizations on social influence. *International Journal of Research in Marketing*, *30*(4), 429–431.

Maniadis, Z., Tufano, F., & List, J. A. (2014). One swallow doesn't make a summer: New evidence on anchoring effects. *The American Economic Review*, *104*(1), 277–290.

Meehl, P. E. (1967). Theory-testing in psychology and physics: A methodological paradox. *Philosophy of Science*, *34*(2), 103–115.

Mishra, A., & Mishra, H. (2011). The influence of price discount versus bonus pack on the preference for virtue and vice foods. *Journal of Marketing Research*, *48*(1), 196–206.

Mogilner, C., & Aaker, J. (2009). "The time vs. money effect": Shifting product attitudes and decisions through personal connection. *Journal of Consumer Research*, *36*(2), 277–291.

Moon, A., & Roeder, S. S. (2014). A secondary replication attempt of stereotype susceptibility (Shih, Pittinsky, & Ambady, 1999). *Social Psychology*, *45*(3), 199–201.

Morgan, N. A., & Rego, L. L. (2006). The value of different customer satisfaction and loyalty metrics in predicting business performance. *Marketing Science*, *25*(5), 426–439.

Mukherjee, P. (2014). How chilling are network externalities? The role of network structure. *International Journal of Research in Marketing*, *31*(4), 452–456.

Müller, H. (2013). The real-exposure effect revisited — How purchase rates vary under pictorial vs. real item presentations when consumers are allowed to use their tactile sense. *International Journal of Research in Marketing*, *30*(3), 304–307.

Müller, H., Lehmann, S., & Sarstedt, M. (2013). The time vs. money effect. A conceptual replication. *International Journal of Research in Marketing*, *30*(2), 199–200.

Müller, H., Schliwa, V., & Lehmann, S. (2014). Prize decoys at work — New experimental evidence for asymmetric dominance effects in choices on prizes in competitions. *International Journal of Research in Marketing*, *31*(4), 457–460.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*, aac4716 DOI:10.1126/science.aac4716.

Orazi, D. C., & Pizzetti, M. (2015). Revisiting fear appeals: A structural re-inquiry of the protection motivation model. *International Journal of Research in Marketing*, *32*(2), 223–225.

Orsingher, C., Valentini, S., & deAngelis, M. (2010). A meta-analysis of satisfaction with complaint handling in services. *Journal of the Academy of Marketing Science*, *38*(2), 169–186.

Pashler, H., & Harris, C. R. (2012). Is the replicability crisis overblown? Three arguments examined. *Perspectives on Psychological Science*, *7*(6), 531–536.

Pashler, H., & Wagenmakers, J. (2012). Editors' introduction to the special section on replicability in psychological science: A crisis of confidence? *Perspectives on Psychological Science*, *7*(6), 528–530.

Peck, J., & Shu, S. B. (2009). The effect of mere touch on perceived ownership. *Journal of Consumer Research*, *36*(3), 434–447.

Reichheld, F. F. (2003). The one number you need to grow. *Harvard Business Review*, *81*, 46–54.

Roney, J. R. (2003). Effects of visual exposure to the opposite sex: Cognitive aspects of mate attraction in human males. *Personality and Social Psychology Bulletin*, *29*(3), 393–404.

Salganik, M. J., Dodds, P. S., & Watts, D. J. (2006). Experimental study of inequality and unpredictability in an artificial cultural market. *Science*, *311*, 854–856.

Schnall, S., Benton, J., & Harvey, S. (2008). With a clean conscience: cleanliness reduces the severity of moral judgments. *Psychological Science*, *19*(12), 1219–1222.

Schooler, J. (2011). Unpublished results hide the decline effect. *Nature*, *470*(7335), 437–439.

Schwarz, N., & Strack, F. (2014). Does merely going through the same moves make for a "direct" replication? Concepts, contexts, and operationalizations. *Social Psychology*, *45*(4), 299–311.

Scott, M. L., Nowlis, S. M., Mandel, N., & Morales, A. C. (2008, October). The effects of reduced food size and package size on the consumption behavior of restrained and unrestrained eaters. *Journal of Consumer Research*, *35*, 391–405.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Boston: Houghton Mifflin.

Shih, M., Pittinsky, T. L., & Ambady, N. (1999). Stereotype susceptibility: Identity salience and shifts in quantitative performance. *Psychological Science*, *10*(1), 80–83.

Please cite this article as: Lynch, J.G., et al., Reflections on the replication corner: In praise of conceptual replications, *IJRM* (2015), http://dx.doi.org/10.1016/j.ijresmar.2015.09.006

Shiv, B., Carmon, Z., & Ariely, D. (2005). Placebo effects of marketing actions: Consumers may get what they pay for. *Journal of Marketing Research*, *42*(4), 383–393.

Shrum, L. J., Lowrey, T. M., Luna, D., Lerman, D. B., & Liu, M. (2012). Sound symbolism effects across languages: Implications for global brand names. *International Journal of Research in Marketing*, *29*(3), 275–279.

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*(11), 1359–1366.

Simons, D. J. (2014). The value of direct replication. *Perspectives on Psychological Science*, *9*(1), 76–80.

Simonsohn, U., Simmons, J. P., & Nelson, L. D. (2014). Anchoring is not a false-positive: Maniadis, Tufano, and List's (2014) 'failure to replicate' is actually entirely consistent with the original. (April 27, 2014). Available at SSRN: http://ssrn.com/abstract=2351926

Simonsohn, U., Simmons, J. P., & Nelson, L. D. (2015). Specification curve: Descriptive and inferential statisics for all plausible specifications. *Presentation at University of Maryland decision processes symposium*.

Simonson, I. (2014, August). Vices and virtues of misguided replications: The case of asymmetric dominance. *Journal of Marketing Research*, *51*, 514–519.

Simonson, I., & Tversky, A. (1992). Choice in context: Tradeoff contrast and extremeness aversion. *Journal of Marketing Research*, *29*(3), 281.

Steenkamp, J. B. E. M., Batra, R., & Alden, D. L. (2003). How perceived brand globalness creates brand value. *Journal of International Business Studies*, *34*, 53–65.

Stroebe, W., & Strack, F. (2014). The alleged crisis and the illusion of exact replication. *Perspectives on Psychological Science*, *9*(1), 59–71.

Van Doorne, J., Leeflang, P. S. H., & Tijs, M. (2013). Satisfaction as a predictor of future performance: A replication. *International Journal of Research in Marketing*, *30*(3), 314–318.

Weun, S., Beatty, S. E., & Jones, M. A. (2004). The impact of service failure severity on service recovery evaluations and post-recovery relationships. *Journal of Services Marketing*, *18*(2), 133–146.

Wright, S. A., Hernandez, J. M. C., Sundar, A., Dinsmore, J., & Kardes, F. R. (2013). If it tastes bad it must be god: Consumer naïve theories and the marketing placebo effect. *International Journal of Research in Marketing*, *30*(2), 197–198.

Yorkston, E., & Menon, G. (2004). A sound idea: Phonetic effects of brand names on consumer judgment. *Journal of Consumer Research*, *31*, 43–51.

Please cite this article as: Lynch, J.G., et al., Reflections on the replication corner: In praise of conceptual replications, *IJRM* (2015), http://dx.doi.org/10.1016/j.ijresmar.2015.09.006