Re-examining the Experiential Advantage in Consumption: A Meta-Analysis and Review

EVAN WEINGARTEN
JOSEPH K. GOODMAN

A wealth of consumer research has proposed an experiential advantage: consumers yield greater happiness from purchasing experiences compared to material possessions. While this research stream has undoubtedly influenced consumer research, few have questioned its limitations, explored moderators, or investigated filedrawer effects. This has left marketing managers, consumers, and researchers questioning the relevance of the experiential advantage. To address these questions, the authors develop a model of consumer happiness and well-being based on psychological needs (i.e., autonomy, relatedness, self-esteem, and meaningfulness), and conduct an experiential advantage meta-analysis to test this model. Collecting 360 effect sizes from 141 studies, the meta-analysis supports the experiential advantage ($d = 0.383$, $95\%$ CI [0.336, 0.430]), of which approximately a third of the effect may be attributable to publication bias. The analysis finds differential effects depending on the type of dependent measure, suggesting that the experiential advantage may be more tied to relatedness than to happiness and willingness to pay. The experiential advantage is reduced for negative experiences, for solitary experiences, for lower socioeconomic status consumers, and when experiences provide a similar level of utilitarian benefits relative to material goods. Finally, results suggest future studies in this literature should use larger sample sizes than current practice.

Keywords: experiential advantage, meta-analysis, well-being, experiences

How can consumers increase their happiness? And how might managers increase the happiness of their consumers to create firm value? Decades of consumer research has made many suggestions: choose from less variety (Iyengar and Lepper 2000), schedule less leisure (Tonietto and Malkoc 2016), avoid explicit compensatory consumption (Rustagi and Shrum 2019), consider how you discuss consumption (Moore 2012), and be more conscientious of time than money (Mogilner 2010), just to name a few. But, one recommendation that has received abundant attention (and hundreds of citations) pertains to experiences: consume more experiences and fewer material goods (Bastos and Brucks 2017; Chan and Mogilner 2017; Gilovich, Kumar, and Jampol 2015; Nicolao, Irwin, and Goodman 2009; Van Boven and Gilovich 2003).

Many researchers propose that purchasing experiences yields greater happiness for consumers than similar material possessions, or what researchers have termed the experiential advantage. This research stream has undoubtedly influenced consumer research and many replications have been published, along with multiple psychological processes (e.g., social value, identity centrality, and memory) to help explain the effect. However, few researchers have questioned the limitations of the effect, compared the

Editor: J. Jeffrey Inman
Associate Editor: Andrew T. Stephen
Advance Access publication 16 September 2020

© The Author(s) 2020. Published by Oxford University Press on behalf of Journal of Consumer Research, Inc. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com • Vol. 47 • 2021 DOI: 10.1093/jcr/ucaa047
relative impact of the different dependent measures used, and/or explored possible moderators and filedrawer effects. This has left managers, consumers, and researchers questioning the broad relevance of the experiential advantage.

Yet, the tradeoff to consumers and managers is real, particularly in today’s modern economy of experiential marketing and materialism (Pieters 2013; Pine and Gilmore 1999; Schmitt 1999). Consumers have access to a wide array of options when it comes to spending their income, and they continually face a choice between purchasing material goods—such as new skis or a new sofa—to experiences—such as a ski vacation or learning to skydive. Similarly, marketing managers must choose whether to invest their limited resources in the product itself (e.g., the brand REI developing new ski equipment), or to the experiences they may offer as well (e.g., the brand REI developing new ski trips for their members).

In this research, we systematically address these questions by developing a broad model of happiness and well-being based on psychological needs (i.e., autonomy, relatedness, self-esteem, and meaningfulness; Baumeister and Leary 1995; Kahneman, Diener, and Schwarz 1999; Ryan and Deci 2000, 2001; Sheldon et al. 2001), and we conduct a meta-analysis on the experiential advantage to test this model. Collecting 360 effect sizes from 141 studies, from both published and unpublished papers, we examine the strength of the effect across a variety of manipulations and dependent measures, examine theoretical moderators of the effect, and test for the impact of publication bias within the literature. While prior investigations into this literature have contributed studies in batches to examine specific theories for why people may prefer experiences to material possessions and for what types of well-being or social outcomes, we quantitatively synthesize a wide set of studies to ascertain if any theoretically relevant moderators may have an impact on when the experiential advantage occurs. We build a broader theory of the experiential advantage as it pertains to theories of happiness, examining how experiences fulfill different psychological needs to contribute to well-being (i.e., relatedness, autonomy, self-esteem, and meaningfulness). Across our dataset, we observe a moderate effect ($d = 0.383$), about a third of which may be attributable to publication bias. We also find differential effects depending on the type of dependent measures used by the researcher, suggesting that the experiential advantage may be more tied to its impact on outcomes about relatedness (i.e., social connection) compared to happiness and willingness-to-pay measures. Moderator analyses support a reduction in the experiential advantage for negative and solitary experiences, for lower socioeconomic status individuals, and when experiences provide a similar level of utilitarian benefits relative to material purchases. Exploratory analyses also suggest studies in this literature should use larger sample sizes than are currently being utilized. We discuss these findings in light of general theories of happiness and their relevance for researchers, consumers, and managers.

THE MATERIAL-EXPERIENTIAL CONTINUUM AND THE EXPERIENTIAL ADVANTAGE

The way consumers perceive options can be critical to their choice of and happiness derived from those options. Prior literature has advanced multiple continua along which items can be perceived to vary. Consumers can pick from expenditures that range from hedonic (e.g., designer clothes) to utilitarian (e.g., a microwave; Chen, Lee, and Yap 2017; Dhar and Wertenbroch 2000) or from vices (e.g., cake) to virtues (e.g., fruit; Shiv and Fedorikhin 1999). Where consumers perceive options to be located on these continua can shift purchase intentions and decisions (Khan and Dhar 2006, 2007). In this research, we concentrate on another continuum that has important implications for consumer theory and well-being: the material-experiential continuum (Carter and Gilovich 2014; Gilovich et al. 2015).

This material-experiential continuum has consequences for consumer happiness and well-being, and the literature suggests an experiential advantage, which posits that experiential purchases lead to greater individual happiness and well-being than comparable material purchases. In a now-classic study, Van Boven and Gilovich (2003, 1194) asked participants to recall an experience or a material possession of at least $100. Participants subsequently reported how happy thinking of the purchase made them, how much the purchase contributed to their life happiness, and whether the purchase was money well spent. Across these and other measures, experiences bested material possessions: people were happier with experiences. Further, the happiness originating from experiences decays much slower than that of material possessions (Nicolao et al. 2009). That is, whereas people may be similarly satisfied initially with their experiences and material possessions, present satisfaction for material possessions drops more sharply than for experiences (Carter and Gilovich 2010). Further, in an attempt to better understand the psychological processes behind the experiential advantage, researchers have replicated the effect and have occasionally examined boundary conditions (Caprariello and Reis 2013; Nicolao et al. 2009).

A MODEL OF THE EXPERIENTIAL ADVANTAGE

The literature exploring the experiential advantage has proposed several psychological processes that can be categorized into three mechanisms as proposed by Van Boven and Gilovich (2003, 1200): experiences have more social value, experiences are more central to one’s identity, and,
due to dynamics of memory and time, experiences better preserve one’s happiness. Many initial investigations into the experiential advantage sought to provide evidence for these mechanisms (Carter and Gilovich 2012; Van Boven et al. 2010). However, these three mechanisms may be composed of multiple sub-categories of psychological need states that are inputs to happiness that experiences fulfill more rigorously than do material possessions (Ryan and Deci 2000, 2001). For example, the social value of experiences reflects both a need for relatedness via conversation with others (Reis et al. 2000) and self-esteem needs for positive self-image (Sheldon et al. 2001). Similarly, the identity-relevance of experiences could reflect component needs of self-esteem (Usborne and Taylor 2010), autonomy, and meaningfulness (Deci and Ryan 2008; Suh 2002).

We decompose these mechanisms into psychological needs based on the happiness and well-being literature (Baumeister et al. 2013; Kahneman et al. 1999; Ryan and Deci 2000, 2001; Sheldon et al. 2001; see figure 1). We build on the existing explanations for the experiential advantage by separating them into how they fulfill different psychological need states, which in turn contribute to well-being. Specifically, we divide measures in the aforementioned three mechanisms into how they feed into these psychological needs, which have consistently been linked to health, happiness, and eudaimonia: autonomy, relatedness, self-esteem, and meaningfulness (Baumeister and Leary 1995; Reis et al. 2000; Sheldon et al. 2001), and it is also a key psychological need of self-determination theory (Ryan and Deci 2001). The need for relatedness, or often referred to as a need to belong, requires meaningful conversations and/or spending time with others in order to experience stronger positive than negative emotions (Mehl et al. 2010; Reis et al. 2000).

Several researchers have argued that experiential purchases are more conducive to fulfilling this relatedness need. Experiences tend to (but not always) entail human interactions and are more enjoyable with others (Caprariello and Reis 2013). Experiential (vs. material) gifts are preferred when the giver and receiver have stronger social connections (i.e., less social distance; Goodman and Lim 2018) and they lead to stronger social connections (Chan and Mogilner 2017). Others have argued that part of the experiential advantage is due to the fact that experiences are more enjoyable to bring up in conversations and are perceived to have higher social currency (i.e., value in conversations; Bastos and Brucks 2017). Experiences therefore provide, consistent with self-determination theory, a way to fulfill needs for relatedness through a means of having conversation-worthy topics with others and a way to feel understood (Howell and Hill 2009). Given the pervasiveness of the importance of social utility from experiences, we hypothesize a positive relationship between experiential (vs. material) purchases and dependent measures that tap into the construct of relatedness.

The Experiential Advantage and Autonomy

The individual desire for autonomy, or feeling like oneself and endorsing one’s own actions, is another of the central psychological needs of self-determination theory and a contributor to well-being (Ryan and Deci 2001). In multiple studies, including those following individuals over time, engagement in activities for more intrinsic reasons or activities that helped people feel like their true selves bolstered daily well-being (Deci and Ryan 2008; Reis et al. 2000; Ryan and Deci 2001; Sheldon, Ryan, and Reis 1996).

There is evidence that purchasing experiences may contribute to these feelings of autonomy. Van Boven and Gilovich (2003, 1200) contend that an individual life is the sum of one’s experiences (however, see Belk 1988 for a different perspective suggesting the important role of material goods and the self). Indeed, multiple experimental studies have shown people believe experiences fulfill the need for autonomy (Howell and Hill 2009; Kim et al. 2016; Thomas and Millar 2013), and that consumers believe experiential purchases reveal more about who they are than their material purchases (Carter and Gilovich 2012). These experiences are more likely to be included in a description of their life story (Carter and Gilovich 2012). If experiential (vs. material) purchases reveal more about a
person’s consumption choices, which reflects their identity, then we would hypothesize a positive relationship between experiential (vs. material) purchases and measures that tap into the satisfaction of autonomy needs.

The Experiential Advantage and Self-Esteem

Another key component of psychological well-being is self-esteem, which is positive evaluations and confidence in the self (Baumeister 1990). Self-esteem is heightened when individuals have clarity in one’s self-concept and cultural identity (Campbell 1990; Usborne and Taylor 2010), and it has been linked to higher levels of happiness and subjective well-being (Sheldon et al. 2001; Usborne and Taylor 2010).

There is also evidence that experiences reveal more about an individual’s identity (Carter and Gilovich 2012) and that material purchases may reflect negatively on one’s feelings of the self (Bastos and Brucks 2017). Notably, senders discussing material possessions may risk their audience drawing a negative inference about them (e.g., a materialistic identity; Van Boven 2005; Van Boven et al. 2010). Whereas collecting experiences may yield positive self-evaluations, concentrating on possessions could impair peoples’ self-esteem (Fournier and Richins 1991). At its worst, such a focus on material goods could lead to vicious cycle of materialism (Pieters 2013). The experiential advantage literature contains several of these types of measures of impressions or evaluations of others (Valsesia and Diehl 2017; Van Boven et al. 2010). If experiences (vs. material purchases) contribute more to one’s positive self-esteem, then we would hypothesize a positive relationship between experiential (vs. material) purchases and measures that tap into an individual’s concern for self-esteem.

The Experiential Advantage and Meaningfulness, Happiness, and Purchase Preference

Another related concept to happiness and subjective well-being is meaningfulness, or eudaimonia, and is a “cognitive and emotional assessment of whether one’s life has purpose and value” (Baumeister et al. 2013, 506). Its roots stem from early theories of self-actualization and it is different from happiness in that it encompasses feelings of meaning and purpose over long periods of time. It is driven by achieving one’s long-term goals or growth (Baumeister et al. 2013; Ryan and Deci 2001).

There is evidence that experiential and material purchases may differ in the extent to which they contribute to feelings of meaningfulness. In the experiential advantage literature, several studies ask about personal growth from the experience or material possession (Sisso 2013), their importance or meaningfulness (Goodman et al. 2016), or about the purchase’s contribution to happiness in life (Nicolao et al. 2009; Van Boven and Gilovich 2003) as one of several happiness-related measures. Further, studies have found activities that contain personal expressiveness (which are linked to experiential purchases) enhance eudaimonia (Waterman 1993). Based on these findings, we hypothesize that experiential (vs. material) purchases will be positively associated with measures of meaningfulness.

Autonomy, relatedness, and self-esteem are also key drivers of meaningfulness and happiness (Ryan and Deci 2000), the latter of which is usually represented as affective measures in the material-experiential literature. Thus, we would hypothesize a positive relationship between experiential (vs. material) purchases and measures of meaningfulness and happiness. Further, given that customers try to maximize their happiness and long-term meaning when making purchase decisions, we would also hypothesize a positive relationship between experiential (vs. material) and measures of purchase likelihood.

Testing the Theoretical Model

We test these psychological processes by coding the various dependent measures used across the 141 studies in our meta-analysis. By separating these measures into separate psychological needs, we can examine the relative effect size strength of each process and compare it to other measures tapping into downstream consequences, namely meaningfulness, happiness, and purchase intentions. Though we cannot test traditional mediation with our data (each study had different dependent measures), we can test whether the direct effect on happiness is larger than the effect on each component of psychological needs.

In addition to testing our theoretical model, we can also test the effect size on purchase intentions relative to happiness and psychological needs. We treat purchase intention measures separately from happiness in this meta-analysis to specifically test whether feelings of happiness actually translate into comparable effects on consumer behavior.

We can also test from this model whether there is evidence for a material-experiential continuum. Other research has proposed that material-experiential difference is a continuum and can be maximized or minimized. For instance, there are manipulations that may minimize this difference through using the same purchase framed as either a material or experience (e.g., a 3D TV or a grill; Bastos and Brucks 2017; Carter and Gilovich 2012). Similarly, there are manipulations that maximize this difference by picking two different purchases at each end of the spectrum (e.g., a new shirt vs. a ski pass; Van Boven et al. 2010, Study 3). Recall tasks (Van Boven and Gilovich 2003) and real purchase tasks (Chan and Mogilner 2017) should fall somewhere between these two extremes because individuals have a wide latitude in what can be invoked (Van Boven and Gilovich 2003). Given the evidence that the material-experiential difference is a continuum, we hypothesize that manipulations that maximize the difference (i.e., two
different purchases), will have larger effect sizes than manipulations that minimize this difference (i.e., same purchase, different framing).

Other Theoretical Factors

The literature has posited other boundary conditions for the experiential advantage. While there are some individual studies that provide evidence, our meta-analysis enables a broader-scale test of the following boundary conditions.

Valence. Some purchases have a positive outcome, while others turn out negatively. Nicolao et al. (2009) report that the experiential advantage only holds for positive purchases, yet it is attenuated for negative cases. We therefore code for whether the experience or material possession were specified to have gone poorly or not, and hypothesize that effect sizes will be smaller for negative (vs. positive) outcomes.

Solitary versus Social. Many experiential purchases are consumed with others (e.g., a concert) while others are consumed alone (e.g., a massage). Caprariello and Reis (2013) provide evidence that the experiential advantage only occurs when the experience is done with others. When an experience is a solitary event, they found no experiential advantage, suggesting we should find a smaller effect size for solitary events. Thus, we code whether the experience was constrained to be a solitary event, and hypothesize a smaller effect size for solitary (vs. social) events.

Utilitarian/Hedonic Value. Material possessions are often described in the literature in terms of their function, whereas experiences are often described more in terms of enjoyable social use (Bastos and Brucks 2017; Carter and Gilovich 2010, 2012; Tully and Sharma 2018). One possibility is that there is a utilitarian advantage for material possessions and a hedonic advantage for experiences. These differences—a positive difference in hedonic benefits for experiences relative to material goods, and a negative difference in utilitarian benefits for experiences relative to material goods—may contribute to the experiential advantage. If so, then we would hypothesize a decrease in the effect size as these differences decrease. Specifically, we expect a smaller effect size when (a) the difference in hedonic benefits between experiences and material goods decreases or (b) the difference in utilitarian benefits between experiences and material goods decreases. Thus, we measure the extent to which the purchases in the literature are considered to be more or less utilitarian and hedonic.

Socioeconomic Status. A consumer’s socioeconomic status may also affect their ability and intention to make hedonic purchases, which may be related to experiential purchases. Lee, Hall, and Wood (2018) provide evidence that the experiential advantage only emerges for higher socioeconomic status individuals. As one illustration, the experiential advantage was stronger for studies with participants from private universities (i.e., higher socioeconomic status) than for studies with participants from public universities (i.e., lower socioeconomic status). We therefore mimic their coding of whether the sample in question is from public or private universities, and hypothesize that effect sizes will be smaller for public (vs. private) universities.

Methodological Factors and Their Impact on the Experiential Advantage

There may also be other methodological factors that moderate the experiential advantage, significantly influencing effect size. While there is not any clear theoretical prediction, these factors may provide insight for future research. Thus, in our meta-analysis, we also examine the following factors.

Publication Year. The year of study publication may be related to effect sizes according to the decline effect (Schooler 2011). Some literatures exhibit changes in effect sizes over time potentially due to methodological refinements (Mooneyham et al. 2012).

Publication Bias. Unpublished studies may have smaller effect sizes because they were not deemed to have strong enough results for publication (Rosenthal 1991; Rosenthal and Rosnow 2008). Alternatively, such studies may have similar effect sizes as the published literature if they explored a potential moderator that did not come to fruition, and the study was not published due to not contributing new insights to extant theory.

Sample. Some samples, such as Amazon Mechanical Turk, can differ from other samples. Amazon Mechanical Turk has been a particularly popular sample in consumer research in recent years, but it has its critics and there are individual differences (e.g., extraversion, depression in some samples) between Amazon Mechanical Turk and...
other samples (Goodman, Cryder, and Cheema 2013; Goodman and Paolucci 2017).

**Dependent Measure Format.** Researchers have used various dependent measures to measure happiness and well-being. Whether the dependent measures were Likert-scales (e.g., happiness questions from Van Boven and Gilovich 2003) or non-scales (e.g., dictator game allocations from Walker, Kumar, and Gilovich 2016) could have an impact on effect sizes.

**Study Design.** Studies may vary in effect sizes based on whether they employ within-subject or between-subject designs. Within-subject designs should technically have higher power than between-subject designs (Rosenthal and Rosnow 2008), which might translate into smaller effect sizes (and tighter confidence intervals) for within-subject designs as an artifact. Thus, in our meta-analysis, we test whether within-subject designs are associated with smaller effect sizes.

**STUDY OVERVIEW**

We conduct a comprehensive and systematic meta-analysis on the literature of the experiential advantage. We examine studies in which the dependent measures can be considered in terms of an objective advantage for experiential purchases versus material purchase (e.g., happiness or feelings of connectedness). We find medium size effects for measures and conditions intended to conceptually replicate the experiential advantage ($d = 0.383$), with the full literature having an overall effect size about a fifth smaller ($d = 0.315$). We rule out some publication bias concerns, and we test key theoretical and methodological moderators of the effect.

**METHODOLOGY**

**Literature Search**

We searched multiple databases and conference programs for studies and papers during Summer 2017. First, we searched both PsycINFO and Proquest Dissertations and Theses using “(experiential OR material OR material-experiential OR experiential-material) AND (gift OR consumption OR product OR products OR purchase OR purchase OR possessions OR possession).” Second, we checked all forward citations of Van Boven and Gilovich (2003) on Web of Science. Third, we looked through programs from the Society of Consumer Psychology (SCP), Association for Consumer Research, Society for Personality and Social Psychology (SPSP), and Society for Judgment and Decision Making (SJDM) conference programs.

We e-mailed authors of papers and studies discovered through the previous step in July 2017 to verify the information from the studies listed and to request missing information. In total, we sent 42 e-mails, of which 36 (85.7%) received responses. A sample e-mail can be found in web appendix A. We also reached out for unpublished materials through e-mails sent to authors of studies and papers, the ACR-L listserv, the Society for Judgment and Decision Making mailing list, and the Society for Consumer Psychology website.

**Inclusion Criteria**

We included studies that adhered to four criteria:

1. **Consequences.** Studies needed to examine downstream consequences of comparisons between experiential purchases (or experiences) and material purchases (or possessions). These consequences typically dealt with well-being (e.g., life satisfaction or happiness; Van Boven and Gilovich 2003; Nicolao et al. 2009), self-esteem (e.g., impressions; Van Boven, Campbell, and Gilovich 2010), or identity-based (e.g., Carter and Gilovich 2012) dependent measures. We enumerate the types of dependent measures we coded in the moderator coding section.

2. **Manipulation.** Studies needed to manipulate experiential versus material purchase either within- or between-subject. A full list of included manipulations can be found in the moderator section.

3. **Advantages.** Dependent measures needed to be interpretable in terms of an advantage. For example, a measure that examined whether experiences led to more happiness with one purchase versus another is an objective advantage. Similarly, whether or not consumers feel more connected with others by an experiential or material purchase we consider to be an objective advantage. However, whether individuals are perceived to be higher or lower on intrinsic or extrinsic motivation (O’Keefe, Caprariello, and Reis 2012, Study 1) or want to consume a good earlier or later (Kumar and Gilovich 2016), or whether individuals experience regret for action or inaction (Rosenzweig and Gilovich 2012) all are not easily translated into advantages without an additional value judgment.

4. **Statistical Sufficiency.** Studies needed to have sufficient information to be able to compute effect sizes. We reached out to authors regarding studies with missing information as part of the literature search. We were unable to acquire sufficient information for one or more studies in the following papers: Ang et al. (2015), Jun (2015), Madan, Lim, and Ng (2016), Millar and Thomas (2009), Min (2012), Nicolao (2009), Yu et al. (2016), and Zhang et al. (2014).

Based on these criteria, we excluded studies that focused on antecedents that influenced preference for material or...
experiential purchases (Tully, Hershfield, and Meyvis 2015), requested participants to recall an unspecified type of good (Guevara and Howell 2015, Study 1; Nicolao et al. 2009, Study 2) or a good that was both experiential and material (Guevara and Howell 2015, Study 1; Sääksjärvi, Helleń, and Desmet 2016, Study 1).

We included studies into two datasets: a Conceptual Replication dataset and a Full dataset. In the Conceptual Replication dataset, we excluded all dependent measures or conditions designed to attenuate or reverse the effect, or data in which the mean effect was predicted to be zero without controlling for other relevant moderators or mediators (Moore 2014; Shim and White 2014). For example, Nicolao et al. (2009) predicted and showed an attenuated experiential advantage for negative recall, so it is excluded. However, this is different than a failed replication that was intended to conceptually replicate the experiential advantage (Park and Sela 2016; Sisso 2013), which were always included. We include measures that serve as alternative explanations in which the authors propose that it is possible experiential versus material may matter (e.g., liking and thoughtfulness from Chan and Mogilner 2017).

In the Full dataset, we include all effect sizes independent of whether they originate from conditions or measures designed to attenuate or reverse the effect. This dataset enables a test of theoretically relevant moderators designed to test boundary conditions of the experiential advantage (e.g., negative valence, Nicolao et al. 2009; solitary experiences, Caprariello and Reis 2013).

Finally, we excluded all dependent measures that were not related to happiness/subjective well-being, purchase intentions, or the candidate psychological needs. Few measures fell into this category; examples include purchase timing importance (Tully and Sharma 2018) or how many options considered (Carter 2009, Study 4B). The codings for the final group of dependent measures can be found in web appendix B, with analysis code in web appendix C.

Meta-Analytic Methodology

We computed effect sizes for all dependent measures in terms of standardized mean differences (Hedges $d$) using means and standard deviations, $t$-tests, $F$ ratios, and log-odds ratios using standard formulas (Borenstein et al. 2009; Johnson and Eagly 2014; Lipsey and Wilson 2001). We adjusted the value by a sample size correction factor as done in other meta-analyses (Borenstein et al. 2009; Hedges and Olkin 1985; Lipsey and Wilson 2001). We used raw, unadjusted means for these calculations without other covariates.

We computed $I^2$ values to assess heterogeneity (Borenstein et al. 2009; Higgins and Thompson 2002; Huedo-Medina et al. 2006). $I^2$ values reflect the heterogeneity present in a sample as a percentage of total variation that is not sampling variation. That is, they are relative, not absolute, measures of whether the heterogeneity in the sample is due to sampling variation (Huedo-Medina et al. 2006). We also provide $\tau^2$ as a measure of absolute heterogeneity.

We calculated the aggregate effect size using a hierarchical linear model (HLM) to address the multiple dependent measures possibly present in some studies that have been used in other research (Bijmolt and Pieters 2001). This model allows us to test for effect size and impact of moderators based on having multiple studies with one or more measures where some measures may have identical moderator values (e.g., year of publication, sample) or different moderator values (e.g., type of measure). Specifically, we employ a three-level model1 as recommended by other research (Van den Noortgate et al. 2013, 2015), in which in the null model, the overall effect size ($d_{jk}$) based on effect sizes $j$, $j$ = 1, . . . , J from study $k$, $k$ = 1, . . . , K is equal to an overall mean ($\gamma_{00}$) plus sampling variation ($\gamma_{j0}$) plus variation between outcomes within a study ($\gamma_{jk}$) plus between-study variance ($\gamma_{0k}$). Or, alternatively, $d_{jk} = \gamma_{00} + \gamma_{j0} + \gamma_{jk} + \gamma_{0k}$.

We run moderators in this model both by themselves (bivariate regressions) and simultaneously (multivariate regressions) to control for the simultaneous impact of related moderators.

Publication Bias

We assessed publication bias using three common methods in the meta-analytic literature: trim-and-fill, PET-PEESE, and $p$-curve (also used to detect p-hacking). We do not use a filedrawer failsafe number, which estimates the number of null-result studies necessary to reduce the effect size to zero (Rosenberg 2005; Rosenthal and Rosnow 2008), due to multiple criticisms such as lacking confidence intervals (Becker 2005).

First, we employed trim-and-fill on funnel plots with contours as one method to verify the extent to which the effect size would change once accounting for unpublished results (Duval and Tweedie 2000a, 2000b; Palmer et al. 2009; Peters et al. 2008). Funnel plots depict effect size against the standard error for a study; larger studies have smaller standard errors (Egger et al. 1997; Light and Pillemer 1984). Trim-and-fill examines the shape of the coordinates on the funnel plot and, based on the mean weighted effect size, fills in the funnel in spaces where studies are potentially missing (Duval and Tweedie 2000a, 2000b). These imputed potentially missing studies are then incorporated into effect size calculations. However, trim-and-fill is sometimes sensitive to heterogeneity in effect

---

1 This includes "sampling variation for each effect size (level one), variation over outcomes within a study (level two), and variation over studies (level three)" (Van den Noortgate et al. 2015); for more discussion, see Scammacca, Roberts, and Stuebing (2014).
sizes (Terrin et al. 2003), so we also test trim-and-fill on the dataset split into sub-divisions. Further, we adopt a modern method of dividing the funnel plot: contours. Funnel plot contours split pieces of the funnel plot into those in which the effect size is significant or not. Contours on funnel plots enable a test of two sources of funnel asymmetries: small-study biases and filedrawer effects (Palmer et al. 2009; Peters et al. 2008). If the missing studies are mostly nonsignificant, then the asymmetry might be due to filedrawer effects: studies that had nonsignificant results were less likely to be published. On the other hand, if the missing studies are from regions of significance, then the asymmetry may be due to variation in small-sample studies that had huge effect sizes paired with small samples to create the funnel asymmetry. To conduct this analysis, we combined effect sizes to the study level (Bediou et al. 2018).

Second, we ran PET-PEESE as a different method to examine what the effect size would be once controlling for publication bias (Stanley 2008; Stanley and Doucouliagos 2014; see also Gervais 2015). PET-PEESE addresses publication bias through a weighted regression of effect size onto either standard error (PET; Precision Effect Test) or variance (standard error squared; PEESE; Precision Effect Estimate with Standard Errors). If the PET estimate is significant, then it is recommended to use PEESE as an estimate of effect size adjusting for publication bias.

Third, we created a p-curve from all published studies included in our final dataset to test whether the underlying studies had evidential value (Simonsohn et al. 2014a, b). Evidential value conveys that selective reporting cannot entirely explain the results (Simonsohn et al. 2014a), which would occur for right-skewed p-curves. On the other hand, it is also possible that published studies are relatively underpowered to detect the effect at hand, in which case there would not be evidential value. It is also possible that the distribution of p-values is flat, which would suggest there is no real effect because a real effect would have more p-values less than .01 (Simonsohn et al. 2014a). Or, it is possible that most p-values are close to p = .05 (left skew) and that there is a high likelihood of p-hacking in the literature (Simmons, Nelson, and Simons 2011). The studies included in the p-curve can be found in web appendix D. Critically, we only include published studies in the p-curve because those are studies for which there is a fixed author declaration of the predictions for each study that is used to determine what inferential test statistic to include in the p-curve.

Moderator Coding

Moderators were coded for all studies and there were no missing values. We separate these moderators into a few categories below. We also standardize moderators (with mean zero and standard deviation one) for the multiple regression.

Manipulation Moderators. These moderators represent different manipulations of experiential versus material purchases. We considered four specific manipulations from the literature: recall, two separate purchases, same purchase different frame, and real purchases. The first manipulation we label recall, and it consists of participants recalling an experiential and/or material purchase from their past (and provided the standard definition of material/experiential, Van Boven and Gilovich 2003, Study 1). This standard recall manipulation appears frequently in the literature, and variations on it also occur occasionally (e.g., positive or negative recall; Nicolao et al. 2009). The second manipulation we label two separate purchases, and it consists of presenting participants with two separate purchases (e.g., a ski weekend vs. a watch; Van Boven et al. 2010). The third manipulation we label same purchase different frame, and it consists of presenting participants with descriptions of the same purchase (e.g., a grill, a 3D television) that are either framed in terms of its experiential qualities or its material qualities by having participants write about those qualities (Bastos and Brucks 2017, Study 1) or consider the experience of enjoying it with friends versus its material position in their home (Mann and Gilovich 2016, Study 4). The last manipulation we label real purchases, and it involves actual purchases with a given budget (Weidman and Dunn 2016, Study 1).

Dependent Variable Moderators. Based on our theoretical development, we coded each dependent measure in the literature (i.e., not the sum or combination of multi-item measures, but each item within the multi-item measure) into six possible dependent measures (±1 if yes, 0 if no): autonomy, relatedness, self-esteem, meaningfulness, happiness, and purchase likelihood. Several measures are arguably representative of multiple categories (e.g., envy has both happiness and relatedness components); thus, we allowed measures to belong to multiple categories to avoid “best guess” types of heuristics forcing measures to belong to one category.

We conceptualize happiness in terms of subjective well-being, which has both an affective (positive affect minus negative affect) and cognitive component (Diener and Lucas 1999), with a focus on the present and distinct from meaningfulness (which is focused on more long-term...
Meaningfulness measures are those that tap into how individuals feel they are meeting long-term goals, achieving growth, or reaching a sense of deeper purpose in life or the universe. Representative measures include satisfaction with life (Sääksjarvi, Hellen, and Desmet 2016), a purchase’s contribution to happiness in life (Walker et al. 2016), or whether the purchase contributed to personal development (Sisso 2013). These also include memory type measures, such as Goodman, Malkoc, and Stephenson’s (2016) measures of memory, and Carter and Gilovich’s (2012) memory exchange measures.

Exploratory Moderators. We code several other factors that are less tied to the theoretical moderators. We code year of study publication, whether a study was published (+1) or not (0) as of the time we conducted our search, whether studies are conducted on Amazon Mechanical Turk (+1) or not (0), whether the dependent measure involves a Likert-scale type question (+1) or not (0), whether the dependent measure is retrospective (+1) or not (0), the number of items on the dependent measure, and whether the study was between-subjects (+1) or within-subjects (0).

Coding. In an initial phase, one author coded the studies, which had their characteristics presented to authors of studies via e-mail during the literature search. An independent set of two research assistants who were blinded to the goals of the research were given the moderators definitions from table 1 and also coded the studies following establishing reliability. The first author then coded relatedness, autonomy, self-esteem, and meaningfulness and established reliability independently with one of these coders. The first author then also coded whether studies manipulated valence (positive or negative), solitary (yes or no), and whether study samples were private universities or public universities (for university samples).

To code hedonic, utilitarian, and memory likelihood/frequency, 226 undergraduates (average age = 20.39, 51.8% female) from a public midwestern university rated 64 items in the literature on four dimensions in a random order; 27 participants were removed for not completing all ratings:

a. (Utilitarian) “When thinking about each of the purchases below, to what extent do you believe it is utilitarian? That is, to what extent do you think it is practical, functional, useful, and/or something that helps you solve a problem?”

b. (Hedonic) When thinking about each of the below, to what extent do you believe it is hedonic? That is, to what extent do you think it is pleasant and fun, enjoyable, and/or something that appeals to the senses?

c. (Likely) Some purchases we remember after a few months, and others we tend to forget to think about. For each purchase below, please think about how likely you would be to think about that
### TABLE 1

<table>
<thead>
<tr>
<th>ID</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authors</td>
<td>Names of authors on paper.</td>
</tr>
<tr>
<td>Paper Title</td>
<td>Title of paper</td>
</tr>
<tr>
<td>Journal</td>
<td>Name of journal in which study was published, or dissertation, or working paper/unpublished.</td>
</tr>
<tr>
<td>Year</td>
<td>Year of publication.</td>
</tr>
<tr>
<td>Published</td>
<td>Whether study is published in a paper in a journal (1 = Yes, 0 = unpublished data, working paper, dissertation, or conference proceedings)</td>
</tr>
<tr>
<td>Study</td>
<td>Study number in paper.</td>
</tr>
<tr>
<td>mTurk</td>
<td>Whether study was run on Amazon Mechanical Turk (mTurk) or not (1 = Yes, 0 = No)</td>
</tr>
<tr>
<td>% female</td>
<td>Percentage of participants who were female.</td>
</tr>
<tr>
<td>Method of Induction</td>
<td>Description of manipulation used in paper for experiential versus material.</td>
</tr>
<tr>
<td>Real_Purchase</td>
<td>Study involved either participants recalling their own goods or dealing with a real good to purchase (1 = Real, 0 = Otherwise)</td>
</tr>
<tr>
<td>Recall_Purchase</td>
<td>Study involved the participant recalling their own example of a material and/or experiential good (1 = Yes, 0 = No)</td>
</tr>
<tr>
<td>PriceRange</td>
<td>For recalled goods, the amount of money set as the floor for the expenditure.</td>
</tr>
<tr>
<td>SamePurchaseDifferentFrame</td>
<td>If the good in question is the same in the experiential and the material condition, but it’s framed as one or the other depending on condition (1 = Yes, 0 = No)</td>
</tr>
<tr>
<td>Two Separate Purchases</td>
<td>Does the study present different goods for experiential and material that are not generated by participants (1 = Yes, 0 = No)</td>
</tr>
<tr>
<td>Subjective Well-being and</td>
<td>Does the DV measure happiness, positive or negative emotions/feelings, regret, whether the money was well spent, and other attitudes about the purchase (1 = Yes, 0 = No)? Note: This does NOT include measures about intentions to buy/purchase or willingness to pay,</td>
</tr>
<tr>
<td>Happiness</td>
<td>Does the DV measure purchase intentions, strength of preference, or willingness to pay/borrow (1 = Yes, 0 = No)?</td>
</tr>
<tr>
<td>PreferenceWTPIntentions</td>
<td>Does the DV measure judgments about social behaviors (1 = Yes, 0 = No)? For instance, this includes questions about talking, word of mouth, likelihood of sharing, likelihood of donating to a cause, willingness to engage in social activities, whether the good helped forge friendships, impressions of others.</td>
</tr>
<tr>
<td>Social</td>
<td>Does the DV measure judgments about the person’s identity or the self (1 = Yes, 0 = No). For instance, this includes questions about insight into the true self, is a part of who they are, allowed them to use their talents</td>
</tr>
<tr>
<td>Identity</td>
<td>Does the DV measure judgments about social behaviors (1 = Yes, 0 = No)? For instance, this includes questions about talking, word of mouth, likelihood of sharing, likelihood of donating to a cause, willingness to engage in social activities, whether the good helped forge friendships, impressions of others.</td>
</tr>
<tr>
<td>Memory</td>
<td>Does the DV measure how memorable something is, how frequently people remember things, whether it’s easy to remember, and related (1 = Yes, 0 = No)?</td>
</tr>
<tr>
<td>Retrospective</td>
<td>Does the DV deal with one of the following: what the participant feels right now, forecasting or guessing something that will happen or would happen in the future, or deals with how they felt in the past? (1 = Retrospective, 0 = No)</td>
</tr>
<tr>
<td>Other</td>
<td>Does the dependent measure(s) not fit into the other categories (1 = Yes, 0 = No)?</td>
</tr>
<tr>
<td>NumberOfItemsScale</td>
<td>Number of items/questions composing a scale for a dependent measure.</td>
</tr>
<tr>
<td>Relatedness</td>
<td>Item involves responding on a given set of scales versus not, an example of the latter being making a choice (1 = Scale, 0 = Otherwise)</td>
</tr>
<tr>
<td>Autonomy</td>
<td>Does the dependent measure involve people rating their perceived or actual connectedness to others or something about engaging with others through conversation, activities, or otherwise? This measure does not include evaluations of others that are not about their relationship.</td>
</tr>
<tr>
<td>Meaningfulness</td>
<td>Does the dependent measure involve whether people are (or feel they are) meeting long-term goals, achieving growth, or reaching a sense of deeper purpose in life or their place in the universe?</td>
</tr>
<tr>
<td>Self-Esteem</td>
<td>Does the dependent measure involve identification or evaluation of positive (or negative) qualities in themselves, evaluations of the self or impressions of (or feelings about) others?</td>
</tr>
<tr>
<td>Valence</td>
<td>Was task about positive or negative purchase.</td>
</tr>
<tr>
<td>Solitary</td>
<td>Was experience intended to be consumed by oneself (solitary) or with others (social).</td>
</tr>
<tr>
<td>University</td>
<td>Was sample from public or private university.</td>
</tr>
</tbody>
</table>

**purchase a few months from now** if you purchased it today.
That is, suppose you purchased it today, and a few months pass. How likely would you be to think about the purchase in a few months?

- (Often) Some purchases we think about a lot, and other purchases we don’t think about very often. For each purchase below, please think about how often you would think about that purchase over the next few months from now if you purchased it today. That is, suppose you purchased it today, and a few months pass. How often would you think about the purchase over the next few months?
The complete results from this survey can be seen in web appendix E. We captured ratings from manipulations of two separate purchases in which we could have comparisons between the material and experiential good (i.e., we calculated the difference in the ratings, as labeled in Table 2); we did not analyze rows using the same purchase framed in one of two ways because such a comparison would have yielded a zero. We could not obtain ratings for the recall manipulations because there were no specified purchases for such manipulations. These ratings yielded 94 observations in the conceptual replication dataset and 107 in the full dataset.

RESULTS

Literature Search

Our search found 360 effect sizes from 141 studies in the Conceptual Replication dataset, and 429 effect sizes from 144 studies in the Full dataset. Table 2 shows a descriptive set of statistics from coded variables for the effect sizes (k). Notably, we received unpublished studies (either under review or otherwise) not otherwise available from conferences, dissertations, or papers from 11 sets of authors.

Overall Effect Size

In the Conceptual Replication dataset, the multilevel model yielded an aggregate effect size of \( d = 0.383 \) (95% CI [0.336, 0.430], \( t = 15.99, p < .001 \)). The \( I^2 \) value is 80.67%, which suggests a higher proportion of true heterogeneity rather than sampling variation behind the effect sizes (Huedo-Medina et al. 2006), noting that the \( \tau^2 \) value (.077) represents the actual amount of heterogeneity. On the other hand, in the Full dataset, the aggregate effect size was \( d = 0.315 \) (95% CI [0.265, 0.365], \( t = 12.48, p < .001; I^2 = 86.16%, \tau^2 = .11 \)). Histograms of effect sizes from the Conceptual Replication and Full datasets are shown in figure 2: there is noticeable variation in the effect sizes sampled in this meta-analysis. Of the effect sizes, 3% are negative, 0.56% are exactly 0, and 14.17% are negative. In the Full dataset, 77.16% are positive, 0.70% are exactly 0, and 22.14% are negative.

Publication Bias

We pursued three methods to address publication bias beyond merely incorporating unpublished data in our analyses: trim-and-fill, PET-PEESE, and a p-curve of the published literature (Simonsohn et al. 2014b).

First, we ran a random-effects trim-and-fill analysis using metaphor (Viechtbauer 2010) to assess the impact of controlling for publication bias on the overall effect size. We first aggregated the effect sizes into one per study, which yielded a similar random-effects effect size of \( d = 0.383 \) (95% CI [0.339, 0.427], \( z = 16.98, p < .001 \)) in the Conceptual Replication dataset. The results of the trim-and-fill analysis are depicted in figure 3A–C. Figure 3A depicts the initial funnel plot of effect sizes plotted against precision (standard error), with a vertical line to mark the aggregate effect size. As discussed earlier, the funnel chart may have a relatively empty bottom left-hand corner due to weaker, smaller studies being less likely to be significant, which would hurt their chances of being published. An inverse-variance weighted regression of effect size onto standard error suggests that there may be an asymmetry in the funnel chart. The adjusted effect size, which inserts 39 studies to rectify the asymmetry in the funnel plot, is \( d = 0.271 \) (95% CI [0.220, 0.322], \( z = 10.48, p < .001 \)). This adjusted effect size is only about a third lower than the original effect size, estimating that less than half of the effect can be attributed to publication bias. In figure 3B, the vertical line has moved to the left, consistent with a smaller effect size once publication bias is addressed. In figure 3C, we center the vertical line around zero and create a contour-enhanced funnel plot (Palmer et al. 2009; Peters et al. 2008). Effect sizes plotted in the middle of the chart are nonsignificant and may otherwise not have escaped the metaphorical filedrawer. Effect sizes outside of the center are significant. That more imputed studies cluster in the nonsignificant part of the funnel plot in figure 3C is consistent with the publication bias notion that nonsignificant studies may not be accounted for.

An alternative theory, in which the asymmetry in the funnel plot is due to the presence of small studies with strong, positive effect sizes, is not supported. That is, fewer imputed effect sizes are on the very left, negative portion of the chart in regions of significance. Similar results can be seen for the Full dataset in figure 3D–F. We ran robustness tests on these trim-and-fill analyses in web appendix F that rule out heterogeneity concerns.

---

3 A known issue in meta-analysis is publication bias; however, efforts to combat publication bias may be subject to concern (Rothstein and Bushman 2012). One such concern is that the authors who reach out for data may have either bias who responds or that the authors disproportionately include their own studies. We address these concerns by avoiding inclusion of the author’s filedrawer beyond published data to prevent skewing the meta-analysis to be a summary of our work. This decision prevents the authors from reaching a predetermined conclusion given their own work.

4 We removed six outliers (five positive, one negative) from the Conceptual Replication dataset found using a Grubbs test (see Zelinsky and Shadish 2018). Excluding these meant the minimum effect size was \( d = -.088 \), and the maximum effect size was \( d = 1.59 \). We similarly removed seven outliers (five positive, two negative) from the Full dataset, leaving a minimum effect size of \( d = -1.25 \) and a maximum effect size of \( d = 1.59 \).
Second, we ran PET-PEESE with the robust variance estimation model to similarly account for what the publication bias-adjusted effect size might be. To recapitulate, we ran a weighted least-squares regression of the effect size onto its standard error (PET) or variance (PEESE). The intercept of these regressions is interpreted to be a publication bias-adjusted effect size estimate. That is, we examined the predicted effect size when the standard error or variance is zero (i.e., the most precise study possible). PET-PEESE’s decision rule is to adopt the PEESE estimate if the PET estimate is significant (i.e., does not have zero in its 95% confidence interval). However, as shown in Table 3, there may be considerable publication bias from the PET estimate in the Conceptual Replication dataset, but there exists an effect in the Full dataset. Again, over a third of the effect may be due to publication bias in PEESE, as shown in Table 3, there may be considerable publication bias from the Conceptual Replication dataset. 

Finally, we present a $p$-curve of the results that exhibits right-skew (Figure 4). Approximately 59% of the $p$-values associated with published studies in the literature are below 0.05; ** 

![Table 2: Meta-Analytic Results](https://academic.oup.com/jcr/article/47/6/855/5906525)
.01, while only 6% of the p-values are between .04 and .05. We observe support for evidential value from both a binomial test examining the proportion of p-values above and below .025 (p < .001), and two continuous tests from the full p-curve (z = −16.56, p < .001) and half p-curve (z = −17.27, p < .001). This half p-curve test examines the distribution of p-values falling below .025. We therefore argue selective reporting may not be able to entirely explain the results in the literature. The aforementioned tests would only become non-significant with the removal of over 18 of the lowest p-values (see figure 5). Given the curve has significant values, we would need to drop almost a third of the relevant p-values. However, even if there is not visible evidence of p-hacking being rampant in the overall literature that does not rule out researchers adding more researcher degrees of freedom to studies (Simmons et al. 2011).

Ultimately, these publication bias analyses suggest that selective reporting and publication bias cannot explain the entirety of the experiential advantage. If anything, we find convergent evidence that the effect size might need to be adjusted downwards by approximately a third to address the impact of publication bias. That is, there are likely studies missing which would otherwise occupy the space of studies with nonsignificant p-values.

Moderators

Table 2 summarizes the results of the moderator analyses and reports effect size estimates for the Conceptual
NOTE:—This analysis depicts the funnel plot of the effect sizes from the Conceptual Replication dataset aggregated to the study level (A, left), the imputed studies as white circles (B, middle), and using contours (C, right). The Full dataset figures are shown in D–F.
TABLE 3

PUBLICATION BIAS RESULTS

<table>
<thead>
<tr>
<th>Test</th>
<th>Conceptual Replication dataset adjusted d (t or z) [p]</th>
<th>Full dataset adjusted d (t or z) [p]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effect size baseline</td>
<td>0.383 (16.98) [&lt;.001]</td>
<td>0.329 (14.19) [&lt;.001]</td>
</tr>
<tr>
<td>Trim-and-fill (imputed)</td>
<td>0.271 (10.48) [&lt;.001]</td>
<td>0.232 (8.72) [&lt;.001]</td>
</tr>
<tr>
<td>Precision Effect Test (PET)</td>
<td>0.075 (1.24) [1.18]</td>
<td>0.140 (2.14) [.034]</td>
</tr>
<tr>
<td>Precision Effect Estimate With</td>
<td>0.229 (6.27) [&lt;.001]</td>
<td>0.245 (6.26) [&lt;.001]</td>
</tr>
</tbody>
</table>

NOTE:—1Uses a random-effects model with one observation per study before imputing studies via trim-and-fill. 2Inserts 39 imputed studies. 3Inserts 31 imputed studies.

WEINGARTEN AND GOODMAN 869

FIGURE 4

P-CURVE OF STUDIES

- Observed p-curve
- Null of no effect
- Null of 95% power
- Tests for right skewness: \( p_{\mu} = .001, p_{\mu} > .001 \)
- Tests for left skewness: \( p_{\mu} = .001, p_{\mu} < .001 \)
- Power estimate: 55% (C[91%, 59%])

NOTE:—The solid line depicts the distribution of observed \( p \)-values, the dotted line maps out what a uniform distribution of \( p \)-values would look like, and the dashed line illustrates a distribution of \( p \)-values against which to test if the studies were powered at 33%.

Replication dataset and the Full dataset. Effect sizes are analyzed using both bivariate and a multivariate models. The bivariate \( t \) reports effect sizes simply for that moderator, not controlling for the other moderators. The multivariate \( t \) reports effect sizes for that moderator when controlling for the manipulation type, dependent measures, and methodological factors (and valence and solitary for the full dataset).

Theoretical Modifiers. We observe three general results. First, in the Full dataset, consistent with the manipulations in the literature (Caprariello and Reis 2013; Nicolao et al. 2009), we observe that there is no experiential advantage in the negative domain \( (d = -0.058, SE = 0.118, t = -0.50, p = .62) \), which is significantly different than the positive domain \( (d = 0.325, SE = 0.025, t = 12.77, p < .001; \text{difference: } t = 3.26, p = .001) \), and that solitary experiences exhibit no experiential advantage \( (d = 0.07, SE = 0.087, t = 0.80, p = .423) \) which is significantly different than social experiences \( (d = 0.322, SE = 0.025, t = 12.76, p < .001; \text{difference: } t = 2.96, p = .003) \).

Second, consistent with Lee et al. (2018), the experiential advantage was stronger for student samples from private universities \( (d = 0.547, SE = 0.050) \) than those from public universities \( (d = 0.309, SE = 0.068, t = 2.83, p = .007) \), suggesting the advantage may be more prevalent for higher socioeconomic status subjects.

Third, we find evidence that experiences have fewer utilitarian benefits \( (M = 2.68) \) compared to material goods \( (M = 4.35) \) in the Conceptual Replication dataset, and as this difference got smaller, so did effect sizes \( (d = 0.089, SE = 0.036, t = 2.46, p = .02) \). In terms of hedonic benefits, experiences had greater hedonic benefits \( (M = 5.19) \) compared to material goods \( (M = 4.22) \); however, this difference did not moderate the experiential advantage smaller \( (d = 0.037, SE = 0.055, t = 0.68, p = .751) \), suggesting that only utilitarian components, and not hedonic, may partly explain the effect.

Theoretical Modifiers—Manipulations. To explore whether variation in effect sizes maps onto manipulations that minimize or maximize the difference between material and experiential, we tested four different manipulations: two different purchases, recall, the same purchase, and real purchases. Although we observed few significant results, we saw results directionally consistent with the mapping of the material-experiential continuum onto consequences. First, two different purchases manipulations (i.e., those that maximized the difference between material and experiential by picking two different purchases) had directionally larger effect sizes \( (d = 0.444, SE = 0.04) \) than same purchase manipulations (i.e., those that likely minimized the difference on the scale; \( d = 0.313, SE = 0.073, t = 1.57, p = .118) \). Also directionally consistent with the continuum, recall manipulations \( (d = 0.373, SE = 0.033) \) had effect sizes between manipulations of two different purchases (recall vs. two different purchases: \( t = 1.34, p = .229 \)).
and the same purchase framed as either a material or experiential purchase (recall vs. same purchase: $t = 0.75$, $p = .453$). All real purchases had directionally smaller effect sizes than the same purchase framed differently ($t = 0.83$, $p = .408$), recall ($t = 1.51$, $p = .134$), and two different purchases ($t = 2.10$, $p = .038$).

**Theoretical Moderators—Dependent Measures.** Next, we tested our theoretical model, which consisted of six components: autonomy, relatedness, self-esteem, meaningfulness, happiness, and purchase intentions (see figure 5).

We found that compared to material purchases, experiential purchases contributed more to each of the four psychological needs (relatedness: $d = 0.558$, SE = 0.044; autonomy: $d = 0.471$, SE = 0.073; self-esteem: $d = 0.277$, SE = 0.057; meaningfulness $d = 0.393$, SE = 0.048). Relatedness showed a significantly larger effect compared to other measures ($t = 4.66, p < .001$). Note that this effect on relatedness, which is defined as a social need, is consistent with our results showing stronger effects when experiences are social (vs. solitary). Autonomy showed a directionally larger effect size compared to other measures ($t = 1.26, p = .208$), whereas self-esteem showed a relatively weaker effect compared to other measures ($t = 2.03, p = .044$).

Consistent with the literature, we also found evidence that, compared to material purchases, experiential purchases led to more happiness ($d = 0.326$, SE = 0.028, $t = 11.76, p < .001$) and greater purchase intentions ($d = 0.354$, SE = 0.065, $t = 5.46, p < .001$). Interestingly, happiness showed a significantly smaller effect compared to the other measures ($t = 3.94, p < .001$).

**Exploratory Moderators.** The meta-analysis allowed us to explore several other potential moderators. Consistent with many publication bias analyses, the unpublished studies had weaker effect sizes ($d = 0.312$, SE = 0.034) than the published studies ($d = 0.446$, SE = 0.032; $t = 2.85, p = .005$). Studies conducted on Amazon’s Mechanical Turk showed smaller effect sizes ($d = 0.341$, SE = 0.035) than those from other samples ($d = 0.422$, SE = 0.033), though the effect was only marginally significant ($t = 1.68, p = .096$). We also found that between-subjects studies ($d = 0.405$, SE = 0.026) had larger effect sizes than within-subjects studies ($d = 0.273$, SE = 0.055; $t = 2.19, p = .03$), which is consistent with the notion that within-subject studies may have higher-power (and thus require smaller effect sizes to be significant compared to between-subject studies due to tighter confidence intervals). However, the number of items in the dependent measure and year of publication had no bearing on effect sizes.

**GENERAL DISCUSSION**

For decades, many theorists have supported the notion of an experiential advantage—that consumers yield greater happiness from purchasing experiences compared to material possessions. Almost two decades of empirical research has tested this hypothesis across hundreds of studies and multiple measures, all with varying degrees of evidence. In this research, we revisited the experiential advantage and tested its limitations through a complete meta-analysis of the published and unpublished literature. Developing a theory of happiness around psychological needs, we re-
conceptualized and re-coded the different dependent measures provided by the literature to compare the relative impact of different theoretical constructs. In addition, we explored possible moderators and publication bias. Across 360 effect sizes from 141 studies, and re-conceptualizing the dependent measures into six sub-constructs, we consistently found a significant experiential advantage \( (d = 0.383, \text{CI} [0.336, 0.430]) \), approximately a third of which can be attributed to publication bias. Our results found significant positive effects of experiential purchases compared to material purchases on multiple psychological need states, with relatedness showing the strongest effect, and general measures of happiness showing the weakest (yet still significant) effect. Supporting this relatedness effect, we found that solitary purchases led to less of an experiential advantage compared to non-solitary, consistent with past theorizations (Caprariello and Reis 2013).

Our results also show that experimental differences can have a significant impact on effect sizes. First, we found that different manipulations had a directional impact on effect sizes, consistent with the theorization that the material-experiential distinction is in fact a continuum. Manipulations with more extreme differences (i.e., two different purchases) on the material-experiential continuum were associated with larger effect sizes, whereas manipulations with smaller differences on the material-experiential continuum (i.e., same purchase framed in both ways) had smaller effect sizes. We also found smaller experiential advantage effects for within-subject designs, scale (vs. non-scale) measures, and Amazon Mechanical Turk samples, the latter of which could be due to income differences or differences in self-esteem with Amazon Mechanical Turk participants (Goodman and Paolacci 2017).

Our results also provide evidence for other important moderators that deserve more attention, namely the role of socioeconomics and how hedonic and utilitarian aspects relate to the material-experiential continuum. Research shows that experiences are less preferred under financial constraints (Tully et al. 2015), suggesting that the ephemeral nature of experiences is devalued for more practical and permanent material purchases (Goodman et al. 2016). While we do find evidence that lower socioeconomic participants may show a smaller experiential advantage, we also find evidence that the experiential advantage decreases when the utilitarian difference between material and experiential purchases decreases.

Finally, this meta-analysis also contributes to a unique space in consumer research to assess how perception of purchases has an influence on a variety of outcomes that are ultimately related to consumer happiness. For instance, meta-analyses have examined how materialism relates to happiness (Dittmar et al. 2014), how considering (e.g., number of options; Chernev, Böckenholt, and Goodman 2015) has an influence on outcomes such as satisfaction, and the linkages between happiness and other outcomes (Lyubomirsky, King, and Diener 2005).

Practitioner Implications

The current research has several implications and insights for practitioners such as managers and policy makers. First, the results suggest that marketers can effectively frame a material purchase as more or less experiential and it will contribute to happiness, albeit to a lesser extent than actually making a purely experiential purchase.

Brands such as REI and Patagonia can (and do) highlight the experiences that their material possessions facilitate, and our results support this strategy. The findings also suggest that the best strategy for framing a purchase as experiential is through highlighting how it satisfies needs for relatedness, where we found our strongest effects. Though, that is not to say that factors of autonomy, self-esteem, and meaningfulness cannot play an important role as well. From a competitive strategy perspective, it may be beneficial to take the strategy not being used by one’s competitor to highlight the experiential aspects of a product. Future work should be conscientious of how to move the experiential advantage to either non-traditional samples or field-like settings that use marketing-based manipulations (see Harrison and List 2004).

It is also important to note that the experiential advantage may be limited to purchases targeting higher income brackets and with some hedonic appeal. Replicating Lee et al. (2018), we observe larger effect sizes for private relative to public universities, the latter of which had more similar effect sizes to Amazon Mechanical Turk. It may be prudent to emphasize experiential aspects of purchases for purchases associated with higher socioeconomic status rather than to blanket apply an experiential framing.

Implications for Running Studies on the Consequences of the Experiential Advantage

Our results have implications for future research exploring the consequences of the experiential advantage. First and foremost, researchers should increase sample sizes to achieve appropriate power. The current studies, excluding archival studies, averaged 148 participants, or 74 per cell, which leads to average power of only .64. To be powered at over .85, the average study would need 124 participants per cell, or 248 total. Of course, this estimate would need adjustment based on whether researchers plan to examine happiness, relatedness, or meaningfulness measures, and whether researchers are aiming to find an interaction (Rosenthal and Rosnow 2008), but the general conclusion remains clear: future studies should increase sample sizes.

We encourage future research employ multiple methods with a variety of stimuli, especially in infrequently represented manipulations such as the real purchases with
laboratory dollars (Chan and Mogilner 2017; Nicolao et al. 2009), to increase generalizability and understand limitations. The most common manipulation in the literature uses recall through the experience sampling method, but others provide exemplar stimuli (e.g., skis vs. a ski trip) or frame the same purchase as more or less experiential. Our results suggest that providing stimuli (Kumar and Gilovich 2015; Van Boven et al. 2010) has a directional boost in effect sizes, but this method may appear the least controlled (even with pretesting) due to price and/or purpose differences between the purchases. Another alternative manipulation is to frame the same purchase as either a material or experiential purchase (Bastos and Brucks 2017; Mann and Gilovich 2016), which may be the most conservative (and require larger sample sizes). However, this may violate the classic definition of an experiential possession as being nontangible (Van Boven and Gilovich 2003).

More broadly, we also encourage more variability in the stimuli, such as different price points and multiple category replicates, for future theory development (see Dhami, Hertwig, and Hoffrage 2004). For example, many studies adopted the $100 price tag from Van Boven and Gilovich (2003; Walker et al. 2016). Sampling a wider set of prices and purchases could reveal a deeper understanding of the operation of the experiential advantage.

Implications for the Broader Theory: What Is Next for the Experiential Advantage?

We recommend a pivot for the experiential advantage literature to establish its own canonical moderators and manipulations. For example, the affect-as-information literature typically has studies employ misattribution or salience manipulations to identify relevant psychological processes (Greifeneder, Bless, and Pham 2011; Schwarz and Clore 2007). Presently, while the experiential advantage literature has had separate investigations into the boundary conditions of the advantage (e.g., valence; Nicolao et al. 2009; solitary use; Caprariello and Reis 2013), there are few unifying moderators that appear in multiple papers. Yet, there is more room for investigation, even for two bigger moderators from the literature—negative valence and solitary experiences. For example, how do people treat mixed-valence experiences relative to material possessions? When are solitary experiences as good as social experiences (see Caprariello and Reis 2013)? And when is a material good viewed as experiential, such as a grill or TV (Bastos and Brucks 2017; Carter and Gilovich 2012) or “experiential products” (Guevarra and Howell 2015)?

We also recommend exploring the role of other psychological needs beyond what we were able to analyze (autonomy, relatedness, self-esteem, and meaningfulness). For example, self-determination theory argues that feelings of competence in one’s abilities is an important input to happiness (Ryan and Deci 2000, 2001; Sheldon et al. 2001), yet competence has received infrequent attention in the experiential advantage literature aside from Howell and Hill (2009). Experiences that also allow consumers to actively practice a skill (e.g., skiing or a guitar lesson) may contribute more to psychological needs of competence than, say, a massage or attending a concert. Alternatively, future studies should examine needs that material possessions fulfill better than experiences, as material possessions have known advantages for signaling (Mandel et al. 2006) and perceived longevity of the purchase (Tully et al. 2015).

Memory, a key component to meaningfulness, also deserves more attention in future research. In contrast to happiness, meaningfulness is about how people think about their past and future over time, and memory is a key component of that process (Baumeister et al. 2013; Goodman et al. 2016). Van Boven and Gilovich (2003, 1200) contend the capacity to concentrate on positive details as time progresses is distinctly true of experiences but less so of material possessions. Further, consumers adapt to material gains faster than experiences (see Frederick and Loewenstein 1999; Nicolao et al. 2009). While there is evidence that consumers overestimate how much they will ruminate over experiences (Tully and Meyvis 2017), there has been limited study into frequency of thinking of experiences and material possessions (Weidman and Dunn 2016) or happiness over time (Carter and Gilovich 2010; Nicolao 2009). More questions about memory dynamics remain. For example, are people more likely to focus on positives (and less on negatives) of experiences over time? Are the negatives as frequent but more salient for material possessions (Johnson, Häubl, and Keinan 2007)?

In terms of meaningfulness, it is important that future work clearly delineates it from happiness, the former of which is more closely related to eudaimonia (Ryan and Deci 2000, 2001; Waterman 1993). Previous work into the experiential advantage has not clearly separated happiness in the moment and well-being in service of long-term meaningfulness.

It is also possible that experiential advantage as we know it has changed with a digital social media age that makes experiential consumption hyper-salient and permanent. A decade ago, experiences were argued to be more evanescent than material possessions given they fade from view and memory rapidly, whereas material possessions (e.g., an iPad) could occupy visible space on the counter for months or years (Carter and Gilovich 2010). The proliferation of social media (e.g., Facebook, Instagram, Twitter) means pictures and videos of experiences can extend longer than the end of a ski trip or beach vacation—perhaps contributing to happiness, or deterring from it. Beyond investigating what people share on social media (Valsesia and Diehl 2017), research should examine how engagement with social media itself may strengthen or weaken the experiential advantage.
Limitations

There are a few limitations to our work that should be noted. First, our model is not exhaustive in terms of all the ways in which experiences may contribute to happiness and meaningfulness. Instead, these are the constructs that have been suggested and measured in the literature, which allow us to test our model.

Second, while most of the studies in our data are experiments, results of meta-analyses are correlational. Although researchers may occasionally run experimental studies alongside meta-analyses to infer causality (see Heath and Chatterjee 1995), the present meta-analysis is less amenable to that possibility. Several of the differences between studies concern populations (e.g., Amazon Mechanical Turk) or nonalignable methodologies, which would mean experimentally manipulating them within the same design (i.e., true random assignment) would be difficult.

Third, while our proposed model suggests mediation, we are not able to test for mediation in our meta-analysis because very few studies measured multiple dependent measures. While there is theoretical support for our proposed model in figure 1, it is possible that there is a direct effect on happiness that does not operate through these factors. Similarly, we should note that the measures for meaningfulness and happiness are very similar conceptually, which makes it difficult to assess mediation as well. Thus, we caution toward overinterpretation and encourage future research into these consequences.

Conclusion

Despite the ubiquity of the experiential advantage and the growth of experiential marketing, consumers continue to buy material goods in large quantities. This pattern begs the question: are consumers mis-forecasting or are there advantages to material goods? Some research has started to investigate this notion, such as when giving gifts (Chun and Hiang 2016; Goodman and Lim 2018), marking a special occasion (Goodman et al. 2016), or delaying a purchase (Goodman et al. 2019; Tully and Sharma 2018), but they do not yet provide a cohesive theory. Moving forward, we hope that researchers even-handedly investigate when material goods may lead to happiness and well-being with the same objectivity as investigating the many benefits of experiences.

DATA COLLECTION INFORMATION

The first author analyzed the data from the meta-analytic database. The second author supervised collection of the data from the laboratory sample in the manuscript in Fall 2019; the first author analyzed these data.

REFERENCES

*References with an asterisk have effect sizes included in the meta-analysis.


*Nicolao, Leonardo (2009), “Happiness, Consumption and Hedonic Adaptation,” unpublished dissertation, Marketing Department, University of Texas-Austin, Austin, TX 3368860.


