When Bigger Is Better (and When It Is Not): Implicit Bias in Numeric Judgments

ELLIE J. KYUNG
MANOJ THOMAS
ARADHNA KRISHNA

Numeric ratings for products can be presented using a bigger-is-better format (1 = bad, 5 = good) or a smaller-is-better format with reversed rating poles (1 = good, 5 = bad). Seven experiments document how implicit memory for the bigger-is-better format—where larger numbers typically connote something is better—can systematically bias consumers' judgments without their awareness. This rating polarity effect is the result of proactive interference from culturally determined numerical associations in implicit memory and results in consumer judgments that are less sensitive to differences in numeric ratings. This is an implicit bias that manifests even when people are mindful and focused on the task and across a range of judgment types (auction bids, visual perception, purchase intent, willingness to pay). Implicating the role of reliance on implicit memory in this interference effect, the rating polarity effect is moderated by (1) cultural norms that define the implicit numerical association, (2) construal mindsets that encourage reliance on implicit memory, and (3) individual propensity to rely on implicit memory. This research identifies a new form of proactive interference for numerical associations, demonstrates how reliance on implicit memory can interfere with explicit memory, and shows how to attenuate such interference.

Keywords: implicit memory, interference, numerical cognition, rating format, mindset, cross-cultural marketing

Ellie J. Kyung (ellie.kyung@tuck.dartmouth.edu) is associate professor of business administration at the Tuck School of Business, Dartmouth College, 100 Tuck Hall, Hanover, NH 03755. Manoj Thomas (manojthomas@cornell.edu) is associate professor of marketing at the Johnson Graduate School of Management, Cornell University, 353 Sage Hall, Ithaca, NY 14853. Aradhna Krishna (aradhna@umich.edu) is the Dwight F. Benton Professor of Marketing at the Ross School of Business, University of Michigan, 701 Tappan Street, Ann Arbor, MI 48109. The authors would like to thank Linda Hagen for her help with German translation; Geoff Gunning, Kimberlee Hayward, and Matthew Paronto for their feedback and editorial assistance; Melissa Li for her research support; and Anne Fries for bringing the *Stiftung Warentest* rating system to their attention. They would also like to thank Mrinal Thomas, Maxine Park, and Roxane Park for their assistance in determining the title of this article through a unanimous vote. Lastly, the authors would also like to thank the reviewers, associate editor, and editor for their invaluable comments, which greatly enhanced the article. Supplemental materials, including experiment stimuli, additional graphs, and supplementary analyses, are available in the web appendix.

Gita Johar served as editor and Rebecca Hamilton served as associate editor for this article.

Advance Access publication December 27, 2016

Cause bigger is better
And big is just the best
If you take my advice
You'll outshine the rest

—Princess Amber's advice on party planning to Sofia (in Disney's *Sofia the First*)

Whenever people are asked to evaluate things they encounter in life—products, job candidates, employee performance, research proposals—numeric ratings are inevitably involved. These numeric ratings can use a format where larger numbers indicate that something is better (1 = bad, 5 = good) or have reversed rating poles where smaller numbers indicate that something is better (1 = good, 5 = bad). The format that people are accustomed to encountering can be culturally determined. For example, in the United States, product ratings and student grade point averages are typically presented with bigger-is-better polarity, and there is a strong association between larger numbers and more positive evaluative judgments, not to

 • Vol. 0 • January 2017
DOI: 10.1093/jcr/ucw079

mention a strong association between "bigger" and "better" that permeates advertising, popular culture, and even children's songs. However, in Germany, it is the opposite—lower numbers indicate higher product quality and better student grade point averages, and people have a culturally determined association between "smaller" and "better." How, then, can varying the polarity of ratings used to evaluate products, students, or job applications influence judgments?

People might find themselves making evaluations based on a rating polarity format they are less used to in cross-cultural contexts (e.g., the German equivalent of *Consumer Reports*, *Stiftung Warentest*, rates products using a smaller-is-better system where 0.5 = very good and 5.5 = unsatisfactory) or simply when an evaluation system uses an opposite rating polarity (e.g., grant applications for the National Institutes of Health in the United States are scored by reviewers on a scale ranging from 1 = exceptional to 9 = poor, and these ratings are used in peer review meetings to make funding decisions). We posit that when people attempt to make evaluations using a rating polarity format opposite to the one they have grown up with, the numerical association they are accustomed to using can bias their judgments because they experience a form of memory interference.

In the memory literature, interference where information from the past inhibits people's ability to use information learned in the future is referred to as proactive interference (Jonides and Nee 2006; Keppel and Underwood 1962; Wickens, Born, and Allen 1963). For example, you might have difficulty remembering a new phone number after a move because the old phone number that you have had for years interferes with your ability to remember the new one. Similarly, after consumers encounter an advertisement for a brand that includes information such as price, attribute, or tagline information, their ability to learn the same information for new brands they encounter later can be impaired due to interference (Blankenship and Whiteley 1941; Burke and Srull 1988; Keller 1987; Keller, Heckler, and Houston 1998; Unnava and Sirdeshmukh 1994).

In our research, we introduce a new form of proactive interference for numerical associations that can systematically bias consumer evaluations. The culturally determined numerical association that people learn over time becomes part of their implicit memory—the type of memory that influences judgment without conscious awareness (Graf and Schacter 1987; Roediger 1990; Schacter 1987). This numerical association in implicit memory can then interfere with people's ability to make evaluations using a newly learned format with opposite rating polarity, resulting in judgments that are less sensitive to numeric differences in quality level. We refer to this new form of proactive interference for numerical associations as the "rating polarity effect."

Over a series of seven experiments, we demonstrate that consumers' decisions are repeatedly and persistently affected by the rating polarity effect, even when consumers are well aware that they are using a system with opposite rating polarity. The numeric association in implicit memory surreptitiously influences consumers' evaluations without their awareness. When asked in an online forum whether the polarity of rating format should influence product evaluations, 79% of participants indicated they thought it should have no effect on their judgments (92 participants, 47% female, M_{age}: 38.7 years). Why is it that people fall prey to this rating polarity effect, yet believe they are immune to it? We hypothesize that this effect stems from interference from implicit memory. We designed experiments to delineate the role of implicit memory in this effect. In doing so, we show that reducing reliance on implicit memory can attenuate the interference effects between implicit and explicit memory. Thus, together the culturally determined numerical association in implicit memory and the rating polarity of the evaluation system determine when "bigger is better," and when it is not.

THEORETICAL BACKGROUND

Implicit Numerical Associations

Decades of research on memory and judgments suggest that two different types of memory processes influence our everyday judgments: information stored in explicit memory, and associations stored in implicit memory (Graf and Schacter 1987; Schacter 1987). Explicit memory is characterized by intentional, conscious recollection of episodic information. In contrast, implicit memory influences a task or judgment without conscious awareness or intent—it encompasses the influence of past exposures and experiences, and can spontaneously, even surreptitiously, influence judgments (Graf and Schacter 1987; Roediger 1990; Schacter 1987).

Previous research suggests that implicit associations can play an important role in numerical evaluations (Adaval and Monroe 2002; Bagchi and Davis 2012; King and Janiszewski 2011; Mishra, Mishra, and Nayakankuppam 2006; Monga and Bagchi 2012; Monroe and Lee 1999; Raghubir and Srivastava 2009; Thomas and Morwitz 2009). These implicit associations can also form based on cultural context. For example, in some cultures people learn to associate numbers with bad luck—13 in the United States; 4 in China, Korea, and Japan; and 7 in Ghana, Kenya, and Singapore (Jahoda 1969; Yates 2007). In others, people with left-to-right reading habits associate larger numbers with a right orientation and smaller numbers with a left orientation, and this spatial-numerical association is weaker for people in cultures such as Iran, where people read from right to left (Dehaene, Bossini, and Giraux 1993).

In this research, we characterize a new type of implicit numerical association: that between numeric ratings of a particular magnitude and an evaluative judgment. For example, people in the United States tend to have an implicit association in memory that bigger is better, naturally associating higher numbers with higher quality, while those in countries such as Germany tend to have an implicit association that smaller is better, naturally associating lower numbers with higher quality. This implicit numerical association influences judgments even in situations where it should not, leading to proactive interference.

Proactive Interference for Numerical Associations

Classic work in proactive interference typically examines how previously learned content in explicit memory interferes with memory for new learned information. For example, in the classic Keppel and Underwood (1962) paradigm, participants were shown a series of nonsense consonant strings (KQF, MHZ, CXJ) one at a time and asked to repeat them back to the experimenter after retrieval times of various lengths. Keppel and Underwood's research showed that the successive recall accuracy for each syllable decreased as a result of proactive interference. Similarly, from work in consumer behavior, when people were exposed to an initial advertisement with information about price or other attributes, and then shown a second advertisement with similar (vs. dissimilar) information, their recall of subsequent information was less accurate (Blankenship and Whiteley 1941; Burke and Srull 1988; Keller 1991).

We examine a new type of proactive interference that can occur between a numerical association in implicit memory and a numerical association explicitly provided in a rating format. Specifically, when using a rating format with polarity opposite to that of the numerical association stored in implicit memory, people's evaluations will be unconsciously shifted in the direction of their implicit numerical association. Thus, when an American consumer who is used to bigger-is-better rating polarity comes across a rating format with smaller-is-better rating polarity, her final evaluation will be unconsciously anchored in the direction of a spontaneously, self-generated implicit numerical association in memory. This anchoring will bias the final evaluation in the opposite direction toward the bigger-is-better rating polarity.

> **H1:** In cultures where people hold the bigger-is-better numerical association in implicit memory, product evaluations will be less responsive to differences in numeric ratings when products are rated using smaller-is-better (vs. bigger-is-better) rating polarity.

When products are rated at multiple quality levels, we expect that the slope of the relationship between quality rating and subjective evaluations will be less steep when the products are rated using smaller- versus bigger-is-better rating polarity.

Note that unlike the conscious anchoring and adjustment process studied by researchers such as Tversky and Kahneman (1974), the anchoring effect in the rating polarity effect is unconscious. Because of this, we posit that the underlying process is more consistent with the literature on subliminal numeric priming (Adaval and Monroe 2002; Mussweiler and Englich 2005) and unconscious stereotyping (Gilbert and Hixon 1991). Specifically, because the anchoring process is unconscious and stems from an implicit association in memory in the rating polarity effect, people do not adjust from an anchor they are not aware they are using (see Mussweiler and Englich 2005 for a more detailed exposition of anchoring effects that are not caused by insufficient adjustments).

The Role of Implicit Memory

Although proactive interference is a complex phenomenon that can be studied from several theoretical perspectives such as memory, learning, automaticity, cognitive control, and metacognitive monitoring, we restrict our focus to unearthing the role of implicit memory because proactive interference for numerical associations is based on an underlying implicit association in memory. If interference from the implicit numerical association, which occurs automatically without awareness, underlies the rating polarity effect, then proactive interference for numerical associations should be moderated by factors that activate or inhibit reliance on implicit memory. We identify three such factors: cultural norms, construal mindset, and individual propensity to rely on implicit memory.

Cultural Norms. Our theorizing suggests that cultural norms are an important moderator of proactive interference in numerical cognition. Thus, in a country such as Germany where people hold an opposite numerical association in implicit memory, we hypothesize that the effect of using a rating with bigger-is-better versus smaller-is-better rating polarity will reverse:

> **H2:** In cultures where people hold the smaller-is-better numerical association in implicit memory, their product evaluations will be less responsive to differences in numeric ratings when products are rated using bigger-is-better (vs. smaller-is-better) rating polarity.

Construal Mindset. Recent research suggests that construal mindset can influence reliance on implicit memory. People make judgments and decisions along a continuum from abstract to concrete (Freitas, Gollwitzer, and Trope 2004; Trope and Liberman 2003; Vallacher and Wegner 1989). In a more abstract mindset, people focus more on higher-level, gist representations in memory, while in a more concrete mindset, people focus more on lower-level, verbatim representations. As it relates to memory, reliance on gist memory increases the use of implicit associations in everyday judgments, whereas reliance on verbatim memory reduces it (Fukakara, Ferguson, and Fujita 2013; Rim, Uleman, and Trope 2009; Smith and Trope 2006). Based on this premise, Fukakara et al. (2013) show that when comparing multi-attribute stimuli, people in an abstract mindset are more likely to rely on gist memory, while those in a concrete mindset rely more on verbatim memory. Using a false recognition paradigm, Smith and Trope (2006; see experiment 4) demonstrate that an abstract mindset increases reliance on implicit associations in gist memory, which in turn increases false recognition. Their results also show that the effect of abstraction on false recognition judgments can be independent of changes in effort or motivation. Similarly, research has shown that people in an abstract mindset are more likely to make spontaneous trait inferences (Rim et al. 2009) and rely on stereotypes when making judgments (McCrea, Wieber, and Myers 2012), both of which rely on implicit associations. Thus, prior research suggests that an abstract mindset increases reliance on implicit memory. Therefore, people should be more susceptible to spontaneous proactive interference from an implicit numerical association under an abstract mindset. A concrete mindset, which reduces reliance on implicit associations in memory, should attenuate this interference.

> **H3:** The rating polarity effect is more likely under conditions of an abstract construal mindset than under conditions of a concrete construal mindset.

Direct Measure of Implicit (vs. Explicit) Memory. People differ in their propensity to rely on implicit associations versus explicit rules in everyday judgments, as can be seen in their performance on dual-task tests that require them to allocate their attention between tasks that require both implicit and explicit memory (De Neys 2006). If the rating polarity effect stems from spontaneous interference from implicit memory when using a particular rating format, it should be more pronounced for people who tend to rely more on their implicit memory, but not necessarily for those who rely more on explicit memory. Based on this premise, we predict:

> **H4:** The rating polarity effect will be stronger for people who chronically rely to a greater extent on implicit memory.

OVERVIEW OF EXPERIMENTAL PARADIGM

Examining the effects of proactive interference requires a specific experimental paradigm. As described by Jacoby (1991), measuring the extent of interference between implicit associations and explicit rules requires comparing outcomes where the implicit association and explicit rule are consistent versus when they are inconsistent. In our experiments we compare two conditions—one where the numerical association of the rating format and the numerical association in implicit memory are consistent (consistent rating polarity), and another where the numerical association of the rating format and the numerical association in implicit memory are inconsistent (inconsistent rating polarity). If there is no interference between the implicit numerical association in memory and the numerical association used in the rating format, then judgments in the consistent and inconsistent rating polarity conditions should be identical. However, if the judgments vary across the two conditions, this is evidence for interference. The difference in evaluations between participants in the consistent versus inconsistent rating polarity conditions reflects the extent of interference (Jacoby 1991).

Note that a key objective of this research is to demonstrate that the rating polarity effect is caused by implicit memory interference and that it manifests without the participants' intention or awareness. To this end, for each experiment we conduct pre-evaluation and post-evaluation comprehension tests. The pre-evaluation comprehension test, conducted after exposure to the rating format but before exposure to the stimuli, ensures that participants have read and understood the information on rating polarity. Participants are allowed to proceed to the experiment only after they pass this test. The post-evaluation comprehension test is done after the experiment and enables us to rule out inattention, miscomprehension, or forgetting as possible alternative explanations for our results.

PRETEST: WILLINGNESS TO PAY

As a preliminary test of this experimental paradigm, we conducted an experiment involving a sealed bid auction. Seventy-two participants at a US university were given the opportunity to bid on a stainless steel mug that was described either as having a quality rating of 6.1 on a scale where 1 = unsatisfactory and 7 = very good (bigger is better) or an equivalent rating of 1.9 on a scale where 1 = very good and 7 = unsatisfactory (smaller is better). Those that read a description using bigger-is-better rating polarity bid significantly more ($M = \$4.42$) than those that read one using smaller-is-better rating polarity ($M = \$2.70$, $F(1, 70) = 4.80$, $p = .03$; see web appendix A for full description). This difference between the two conditions offers preliminary support for the proposed interference effect.

EXPERIMENT 1: VISUAL PERCEPTION

Experiment 1 examines the rating polarity effect in the domain of visual perception. Can the rating polarity effect distort visual perception and fool our eyes? All participants were shown the same set of before and after photographs, and we examined whether the rating polarity effect influenced their visual perception across products of both high and low quality. Comprehension, mood, and need for cognition were also measured to rule out potential alternative explanations.

Method

Participants. One hundred seventy-two undergraduates from a US university (63% female; average age: 21.1 years) participated in this computer study in exchange for course credit. They were randomly assigned to one of four conditions in a 2 (rating polarity: consistent vs. inconsistent) × 2 (quality level: low vs. high) between-subjects design.

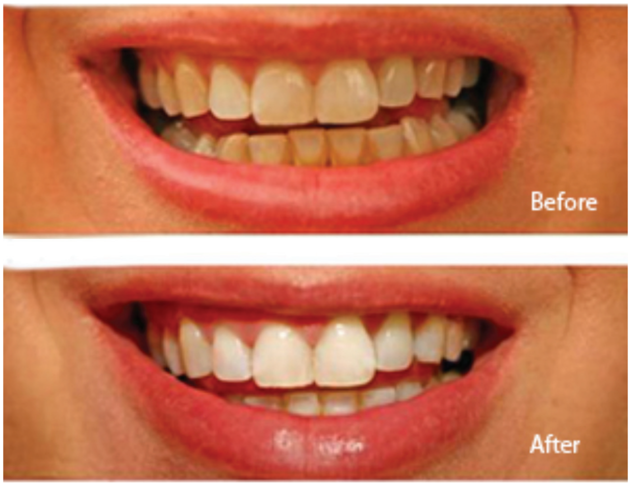
Procedure. Participants were told they would be presented information about a new peroxide-free, home teeth-whitening product currently available in Europe that the producer is considering launching in the United States. Their task was to evaluate how much change they saw in the before versus after photographs accompanying the product information. Participants were told that a quality rating would be provided by a reputable consumer welfare agency in Europe known for its evaluation of consumer products. Those in the consistent rating polarity condition were told that 1 = unsatisfactory and 7 = very good, while those in the inconsistent rating polarity condition were told that 1 = very good and 7 = unsatisfactory. Half the participants were given a high quality rating (consistent condition: 6.1; inconsistent condition: 1.9) and half were given a low quality rating (consistent condition: 1.9; inconsistent condition: 6.1). Participants were asked to confirm that they understood this rating format [Yes/No] before proceeding.

Pre-evaluation Comprehension Test. Participants were then asked the meaning of the 1 and 7 rating poles for the quality ratings [very good or unsatisfactory]. If they responded incorrectly, a message asked them to correct their response, and they could proceed only after answering correctly. This was done to ensure that the results did not stem from inattention or miscomprehension of the rating poles.

Visual Perception. All participants were then shown the same photographs of teeth before and after treatment (see figure 1) and descriptive information, which included the quality level for the product. The before and after photographs were pretested as showing moderate improvement. The quality level rating that participants were shown was either 6.1 or 1.9; each quality rating corresponded with a low- versus high-quality product depending on the rating polarity. Participants were then asked: "How much whiter do the teeth look in the 'after' versus the 'before' photograph?" [not at all whiter/much whiter], "How much cleaner do the teeth look in the 'after' versus the 'before' photograph?" [not at all cleaner/much cleaner], and "How much improvement do you see in how the teeth look after using the whitener?" [no improvement/significant improvement]. These questions were asked using an unnumbered slider scale, but coded as 0 to 100 by the program so as to avoid any numerical association (see web appendix B). The three measures were used as an index of visual perception.

FIGURE 1

EXPERIMENT 1 STIMULI: BEFORE AND AFTER PHOTOGRAPHS



Post-Evaluation Comprehension Test. After entering their bid, participants were asked the meaning of the rating poles 1 and 7 [very good or unsatisfactory]. This was done to confirm that the participants did not become confused or forget about the meaning of the ratings.

Rating Polarity Typicality. Participants were also asked to indicate which of the two numerical associations is more typical: "Higher numbers indicate better quality" or "Lower numbers indicate better quality." They could also choose a third option labeled "Not sure."

Additional Measures. To rule out possible alternative explanations, participants were also asked questions about their current mood [slider scales anchored at bad/good, unpleasant/pleasant, negative/positive], given the 18 items from the short-form Need For Cognition (NFC) scale (Cacciopo, Petty, and Kao 1984), and asked demographic questions.

Results and Discussion

Post-Evaluation Comprehension Test. After the main task, all participants answered the two-rating-pole comprehension test questions. Two percent of participants in the inconsistent and 0% of participants in the consistent rating polarity conditions did not answer both of these questions correctly. Participants who answered either of these questions incorrectly were excluded from the analysis for this and all subsequent studies, although including them did not change any of the main results. We followed this procedure in each of our experiments to ensure that the reported results cannot be ascribed to inattention or miscomprehension.

Rating Polarity Typicality Check. A majority of participants (83.7%) indicated that the bigger-is-better numerical association is more typical, 14.0% indicated a smaller-is-better association is more typical, and 2.3% indicated they were not sure, confirming the bigger-is-better numerical association is more dominant in implicit memory for US consumers.

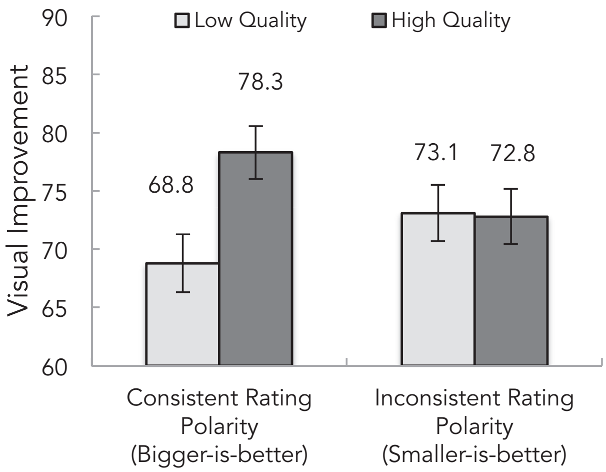
Visual Perception. A two-way ANOVA with rating polarity and quality level as independent measures and the visual perception index as the dependent measure (α = .82) revealed a marginally significant effect of quality level. Participants saw the higher-quality product as more effective ($M = 75.5$) than the lower-quality product ($M =$ 70.9, $F(1, 165) = 3.71$, $p = .06$). More importantly, we obtained a significant interaction ($F(1, 165) = 4.17$, $p = .04$). Planned comparison tests revealed, as predicted, that in the consistent rating polarity conditions (bigger is better), participants saw the high-quality product as demonstrating significantly greater improvement in the before versus after photographs ($M_{high} = 78.3$) than the low-quality product ($M_{low} = 68.8$, $F(1, 165) = 7.48$, $p < .01$). However, in the inconsistent rating polarity conditions (smaller is better), the effect of quality rating is attenuated ($M_{high} = 72.8$ vs. $M_{low} = 73.1$, $F(1, 165) = .007$, $p = .93$; see figure 2). When the rating polarity used a numerical association inconsistent with the numerical association in implicit memory, participants' evaluations of visual improvement were less responsive to differences in quality level, supporting hypothesis 1. Web appendix C summarizes the means and 95% confidence intervals by conditions across all experiments, and for further data visualization, also plots the means and 95% confidence intervals for the consistent versus inconsistent rating polarity conditions.

Ruling Out Alternative Explanations (Mood and NFC). The three measures of mood were averaged into an index ($\alpha = .93$), and a two-way ANOVA with rating polarity and quality level as independent measures revealed no significant main effects (ps > .80) or interaction (p = .13) on mood. Thus, negative mood from a format with inconsistent rating polarity cannot account for the results.

We also conducted a linear regression analysis where the independent measures were rating polarity (dummy-coded: consistent = 0, inconsistent = 1), the mean-centered NFC score, and the interaction of the two with visual perception as the dependent measure. We obtain no significant main or interaction effects for NFC (ps > .57),

FIGURE 2

EXPERIMENT 1: EFFECT OF QUALITY LEVEL ON VISUAL PERCEPTION BY RATING POLARITY (US)



NOTE.—All errors bars represent standard errors.

indicating that the rating polarity effect is independent of NFC, suggesting that even participants with high need for cognition are susceptible to the rating polarity effect and may not be aware of its effects.

Experiment 1 provides supporting evidence for the rating polarity effect: participants perceived the visual change in before and after photos as less impressive when the quality rating was reported using smaller-is-better polarity, even though they all viewed the exact same set of photos. Their judgments were less responsive to differences in product quality when using a rating system with polarity inconsistent with the numerical association they hold in implicit memory. Confusion, forgetting, differences in need for cognition, and mood cannot account for the results.

Although this experiment and the pretest provide evidence of the rating polarity effect, there are several important questions that remain. First, does the rating polarity effect manifest for only single judgments, or will it continue to manifest across repeated judgments? If the rating polarity effect stems from spontaneous interference from implicit memory without people's awareness, it should continue to manifest over multiple judgments, even as participants have more practice using a format with inconsistent rating polarity. Second, does the rating polarity effect persist for multiple levels of product quality? And third, does the rating polarity effect occur only with particular types of response alternatives? If it stems from spontaneous interference from implicit memory, it should be robust to multiple types of response alternatives. The following two experiments address these questions.

EXPERIMENTS 2A AND 2B: REPEATED EVALUATIONS

Experiments 2a and 2b test for the rating polarity effect in a more conservative context where each participant evaluates 15 products across five levels of product quality. If the rating polarity effect is caused by confusion or inattention, then it should not manifest in a repeated-measures design where each participant has the opportunity to evaluate many products and learn from experience. However, if it stems from spontaneous proactive interference from implicit associations, the effect should manifest even with a repeated-measures format, further supporting hypothesis 1. In addition, to further rule out the possibility that the rating polarity effect manifests due to orientation of the response format, we ran the experiments using two different response formats: purchase intent as a binary yes/no measure (experiment 2a) and willingness to pay as a drop-down list of dollar amounts in 50-cent increments (experiment 2b). Since the two experiments were identical in procedure and differed only in the response format, we report the experiments together. Because differences in mood and need for cognition did not affect the results of the first experiment or any subsequent experiments, these analyses for possible alternative accounts are detailed for this and all subsequent experiments in web appendix D.

Method for Experiments 2a and 2b

Participants. U.S. based participants on MTurk (verified by IP address) participated in the experiments in return for $1.50: 213 participants in experiment 2a (49% female; average age: 37.1 years) and 221 participants in experiment 2b (50% female; average age: 36.3 years). They were randomly assigned to the consistent or inconsistent rating polarity condition within each experiment.

Procedure. Participants were told that the study was being conducted by a large retail store to understand American consumers' evaluations of European brands that might be introduced in the United States. They were shown brands of five quality levels (1, 2, 3, 4, and 5) in each of the three different product categories (water, margarine, and toothpaste). Presentation order was randomized for each participant across the 15 brands (see web appendix B for example stimuli). For each brand, participants saw the brand name, the quality level (ostensibly taken from *Consumer Reports*), a photo, and a short tagline communicating the brand positioning.

Participants assigned to the consistent rating polarity condition (bigger-is-better) were informed that 1 = inadequate, 2 = adequate, 3 = fair, 4 = good, 5 = very good. Those in the inconsistent rating polarity condition (smaller-is-better) were informed that 1 = very good, 2 = good, 3 = fair, 4 = adequate, 5 = inadequate.

Pre-Evaluation Comprehension Test. Before proceeding to the main task, participants were asked five pre-evaluation comprehension test questions: "If a product had a quality rating of X, what does it mean?" where X was 5, 4, 3, 2, or 1, respectively. They were asked to select one meaning for each rating value from the options very good, good, fair, adequate, and inadequate. If a response to any question was incorrect, a message appeared informing the participant which responses were incorrect and asking them to correct those answers. Participants could proceed to the main task only after correctly answering all five questions.

Key Dependent Measure. For purchase intent (the key dependent measure in experiment 2a), participants were asked, "Would you purchase this product" [Yes/No]. For willingness to pay (the key dependent measure in experiment 2b), participants were asked to indicate their willingness to pay for each brand in US dollars using a drop-down list with price options increasing in 50-cent increments from $0 to $6.00.

Post-Evaluation Comprehension Test. After evaluating all 15 brands, participants were again asked to indicate the meaning of the rating poles 1 and 5 using the same response options as in the pre-evaluation comprehension task.

Rating Polarity Typicality. Participants were asked the same typicality question as in experiment 1.

The experiment ended with demographic questions—age, gender, marital status, measures of mood (seven-point scales anchored at bad/good, unpleasant/pleasant, negative/positive) and involvement (five-point scales anchored at not at all/very).

Results and Discussion

Post-Evaluation Comprehension Test. Participants who responded to at least one of the rating pole comprehension test questions incorrectly were excluded from the analysis. In experiment 2a, this included 0.5% of the participants in the consistent and 4.7% of participants in the inconsistent rating polarity condition. In experiment 2b, this included 0.9% of the participants in the consistent and 0.5% of participants in the inconsistent rating polarity condition.

Rating Polarity Typicality Check. A vast majority of participants indicated that the bigger-is-better numerical association is more typical (2a: 97%; 2b: 96%) than the smaller-is-better one (2a: 1.5%; 2b: 3%), while a small percentage were not sure (2a: 1.5%; 2b: 1%), confirming that the bigger-is-better numerical association is more dominant in implicit memory.

Purchase Intent (Experiment 2a). Our objective was to test whether rating polarity moderated the effect of quality level on purchase intent (yes = 1, no = 0). To do so, we dummy-coded rating polarity (consistent = 0, inconsistent = 1), and coded quality level as a continuous variable (inadequate = 1 to very good = 5). Quality levels in the inconsistent rating polarity condition were reverse-coded so that higher numeric scores also indicated higher quality level.

Purchase intent was submitted to a repeated-measures logistic regression analysis using PROC GENMOD in SAS with three independent variables: quality level, rating polarity, and their interaction. This method treats quality level as a repeated, within-subjects continuous variable, and rating polarity as a between-subjects categorical variable. Because category effects do not affect the main results across our studies, we included category (toothpaste, margarine, and water) as an additional within-subjects factor but report results aggregated across the three product categories across all experiments.

The simple effect of rating polarity was marginally significant ($\beta = .76$, $p < .06$), while the simple effect of quality level was significant ($\beta = 1.47$, $p < .01$). Most importantly, the two-way rating polarity × quality level interaction was significant ($\beta = -.34$, $p < .01$). The signs of the coefficients suggest that a greater proportion of participants are willing to purchase brands with higher quality levels, but the effect of quality level is weaker for participants evaluating brands rated with a format using inconsistent versus consistent rating polarity, supporting hypothesis 1. To corroborate that these results were not an artifact of the regression assumptions, we plotted the percentage of participants who were willing to purchase the products for the two rating polarity conditions for each level of quality in figure 3. The pattern of means is consistent with our interpretation of the results.

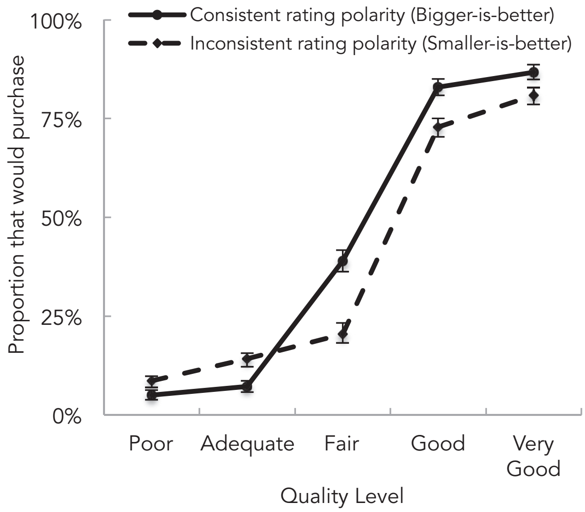
Note that the coefficient for the rating polarity × quality level interaction term represents the extent of interference that stems from the rating polarity effect. Statistically, it is the difference in evaluations when using a rating format that employs rating polarity that is consistent (bigger-is-better) versus inconsistent (smaller-is-better) with the numerical association in implicit memory.

Willingness to Pay (Experiment 2b). We used the same coding for independent variables and analysis procedures as experiment 2a, but submitted the willingness-to-pay measure to a repeated-measures regression analysis using PROC MIXED in SAS.

The simple effect of rating polarity was significant ($\beta = .35$, $p < .01$), as was the simple effect of quality level ($\beta = .54$, $p < .01$). The two simple effects were qualified by a significant two-way rating polarity × quality level interaction ($\beta = -.10$, $p < .01$). Supporting hypothesis 1 Freitas et al. (2004), the signs of the coefficients suggest that willingness to pay increases with higher quality levels, but the effect of quality level is weaker for participants evaluating brands rated with a format using

FIGURE 3

EXPERIMENT 2A: EFFECT OF QUALITY LEVEL ON PURCHASE BY RATING POLARITY (UNITED STATES)



inconsistent versus consistent rating polarity. When participants used a format with rating polarity consistent with their implicit numerical association, on average, a one-level change in quality level changed willingness to pay by \$0.54. However, when the rating polarity was inconsistent with their implicit numerical association, participants' willingness to pay changed by \$0.44 points for a one-level change in quality.

Experiments 2a and 2b provide further evidence of the robustness of the rating polarity effect across 15 repeated measures and five levels of product quality. Product evaluations, measured through binary intent to purchase and vertically oriented willingness to pay, are less responsive to differences in quality level when products are rated using a format with a rating polarity that is inconsistent (vs. consistent) with the implicit numerical association in memory. Mood and involvement cannot account for the results (web appendix D).

THE ROLE OF IMPLICIT MEMORY IN THE RATING POLARITY EFFECT

The first three experiments provide evidence for the rating polarity effect and its robustness, even across repeated judgments and a variety of dependent measures with both horizontal and vertical orientation. But what causes the rating polarity effect? One question that might come to mind is whether the rating polarity effect is caused by misattribution of the subjective experience of fluency. Conceivably, judgments based on formats with consistent rating polarity are more fluent or easier to make than those with inconsistent rating polarity. Preference fluency is usually associated with more favorable evaluations (Alter and Oppenheimer 2009; Lee and Labroo 2004; Schwarz 2004). The observed pattern of results does not support the fluency account because we do not observe only an effect of rating polarity where using inconsistent rating polarity results in more unfavorable evaluations overall. Instead, we observe an interaction effect such that using inconsistent rating polarity results in less sensitivity to differences in product quality overall.

Furthermore, task disfluency or judgment difficulty is typically associated with increased response latency. To test whether the observed effects can be ascribed to difficulty, we conducted supplemental analyses using reaction time as a covariate (see web appendix E). Even when we include reaction time as a covariate, the effect of rating polarity on judgments remains significant. Thus the observed effects cannot be ascribed to judgment difficulty.

We propose that the rating polarity effect is caused by spontaneous interference from implicit memory. To support our proposition that the interference from implicit memory is spontaneous and occurs without the awareness of the influence of this interference on judgments, note that we include only those participants that correctly identify the rating polarity used for their judgments at the end of the judgment task. Yet even participants with a high need for cognition are unable to overcome the effect of rating polarity (see web appendix D). Similarly, participants are unable to overcome the rating polarity effect over 15 repeated judgments. We conducted additional analyses including order as an independent variable for experiments 2a, 2b, and all of the experiments that follow (see web appendix F). In all of these experiments, the two-way interaction representing the rating polarity effect remains significant, while the three-way interaction between rating polarity, quality level, and order is not significant in any of the experiments, indicating that the rating polarity effect does not dissipate over 15 judgments.

If the rating polarity effect stems from consumers automatically relying on implicit memory when making judgments, then it should be moderated by factors that influence the implicit numerical association itself or the tendency to rely on implicit memory when making judgments, such as cultural norms that influence numerical associations held in implicit memory (experiment 3); mindsets that increase reliance on implicit versus explicit memory (experiments 4a and 4b); and measuring direct reliance on implicit memory (experiment 5).

EXPERIMENT 3: MODERATION BY CULTURAL NORMS

In a country where a smaller-is-better numerical association is more likely to be held in implicit memory, such as Germany, using a rating format with bigger-is-better rating polarity should result in product evaluations that are less responsive to differences in quality level, supporting hypothesis 2. Experiment 3 tests this proposition with German participants.

Method

Participants. One hundred members of an online panel in Germany (51% female; average age: 43.0 years) participated in this experiment in exchange for approximately €4.

Procedure. The stimuli were nearly identical to experiments 2a and 2b, with the following differences. The questionnaire was translated into German, and the cover story was slightly modified for a German audience. Participants were informed that a retail store is considering introducing several European brands in the US market and that the quality ratings were provided by a "reputable consumer welfare protection agency widely respected in the US for its unbiased evaluation of consumer products." To test the robustness of the previous results, we used a seven-point, non-numeric semantic differential scale of purchase intentions (anchored at unlikely to buy and likely to buy, with the center labeled as neutral) as the main dependent variable in this study and the studies that follow.

Results and Discussion

Post-Evaluation Comprehension Test. Twelve percent of participants in the inconsistent and 10% in the consistent condition did not correctly answer both rating-pole comprehension test questions. As in the previous experiments, these participants were excluded from the analysis, so the results could not be ascribed to confusion or forgetting (but all effects persist when we include these participants).

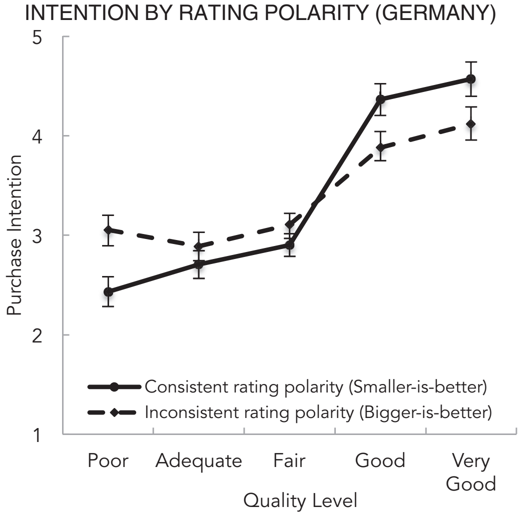
Rating Polarity Typicality. Forty-nine percent of the participants indicated that the smaller-is-better numerical association was more typical, 31% indicated that they found the bigger-is-better numerical association more typical, and 20% indicated they were not sure. These results suggest that although the smaller-is-better format is more typical in Germany, it is not as ubiquitous as the bigger-is-better format is in the United States.

Purchase Intentions. We used a regression model identical to experiment 2b, except that in this case, the smaller-is-better rating polarity was the consistent rating polarity condition and the bigger-is-better rating polarity was the inconsistent rating polarity condition. Purchase intention was submitted to a repeated-measures regression analysis with quality level, rating polarity, and their interaction as predictors. The simple effect of rating polarity was significant ($\beta = .86$, $p < .01$), as was the simple effect of quality level ($\beta = .59$, $p < .01$), but most importantly, the two-way rating polarity $\times$ quality level interaction representing interference from the rating polarity effect was significant ($\beta = -.28$, $p < .01$).

Mirroring the results of experiments 2a and 2b, purchase intention increases with higher quality levels, but consumer evaluations are less responsive to changes in quality level when the brands were rated using a format with rating polarity that is inconsistent versus consistent with the numerical association in implicit memory. However, in this case, it is the bigger-is-better rating polarity that is inconsistent with the implicit numerical association in memory. When participants evaluate brands using rating polarity consistent with their implicit numerical association (smaller-is-better), on average, a one-level difference in quality level changes purchase intention by 0.59 points; however, it changes purchase intention by only 0.31 points when participants use rating polarity inconsistent with their implicit numerical association (bigger-is-better), supporting hypothesis 2. The means of purchase intention are plotted in figure 4.

FIGURE 4

EXPERIMENT 3: EFFECT OF QUALITY LEVEL ON PURCHASE INTENTION BY RATING POLARITY (GERMANY)



Moderation by Implicit Numerical Association. Among German participants, approximately half of the participants held the smaller-is-better numerical association as more typical (49%). However, the remaining participants (51%) did not consider smaller-is-better more typical. If the rating polarity effect stems from spontaneous proactive interference from the implicit numerical association in memory, then it should be moderated by the strength or flexibility of this implicit numerical association. More specifically, those participants who have the implicit numerical association that smaller numbers indicate higher quality should continue to exhibit the rating polarity effect. But those participants who hold a more flexible implicit numerical association (e.g., are not sure about the association or find a bigger-is-better association more typical in a country that typically uses smaller-is-better rating polarity) are less likely to experience as strong of an interference effect from a competing smaller-is-better implicit numerical association when using a rating system with opposite rating polarity. Thus, they are less likely to exhibit the rating polarity effect.

To test this prediction, we submitted purchase intentions to a repeated-measures regression analysis using PROC MIXED in SAS with implicit numerical association (bigger is better or not sure = 0, smaller is better = 1) as an independent variable in addition to rating polarity and quality level, and their two- and three-way interaction terms as predicting variables (see table 1). The three-way interaction between rating polarity, quality level, and implicit numerical association was significant ($\beta = .56$, $p < .01$). Follow-up contrasts reveal that for those participants with a smaller-is-better implicit numerical association, the rating polarity × quality level interaction is significant ($\beta = -.54$,

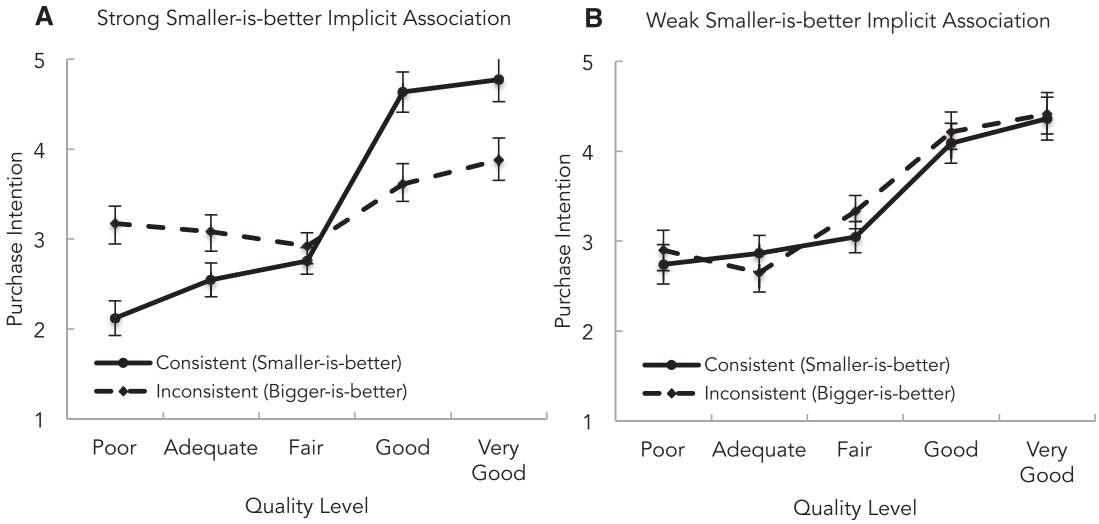
TABLE 1

EXPERIMENT 3 REGRESSION RESULTS: MODERATION BY IMPLICIT NUMERICAL ASSOCIATION

	β	SE	DF	*t*	*p*
Intercept	1.15	0.27	85	4.19	<0.01
Rating polarity	1.60	0.38	85	4.26	<0.01
Implicit numerical association	0.93	0.39	85	2.40	0.02
Rating polarity × implicit Numerical association	–1.56	0.55	85	–2.85	0.01
Quality level	0.74	0.06	1242	13.23	<0.01
Rating polarity × quality level	–0.54	0.08	1242	–7.11	<0.01
Implicit numerical association × quality level	–0.29	0.08	1242	–3.70	<0.01
Three-way interaction	0.56	0.11	1242	5.00	<0.01
PLANNED CONTRASTS					
Implicit numerical association is smaller is better					
Rating polarity × quality level	–0.54	0.08	1242	–7.11	<0.01
Implicit numerical association is bigger is better or not sure					
Rating polarity × quality level	0.01	0.08	1242	0.16	0.87

FIGURE 5

EXPERIMENT 3: MODERATION BY IMPLICIT NUMERICAL ASSOCIATION (GERMANY)



$p < .01$), but for those participants either with a bigger-is-better implicit numerical association or unsure of their numerical association, the rating polarity effect does not manifest ($\beta = .01$, $p = .87$; see figure 5 for pattern of means).

Experiment 3 provides initial support implicating the role of implicit memory in the rating polarity effect. Culture moderates the rating polarity effect: in Germany, where the smaller-is-better numerical association is more common, the rating format with bigger-is-better polarity results in more muted responses to differences in quality level, supporting hypothesis 2. Furthermore, the additional regression analysis with implicit numerical association as a moderator revealed that the rating polarity effect is eliminated for those participants with a more flexible implicit numerical association in memory. This suggests that people who have a less strongly held implicit numerical association in memory do not experience the same spontaneous interference from a numerical association. Mood and involvement cannot account for these effects (web appendix D).

While experiment 3 provides evidence that having different implicit numerical associations can reverse the effect of using ratings of different polarity and attenuate the rating polarity effect, it does not specifically implicate the reliance on implicit versus explicit memory. The remaining experiments directly examine the role of reliance on implicit memory.

EXPERIMENT 4A: MODERATION BY PRIMED CONSTRUAL MINDSET

As discussed earlier, previous research has shown that people in an abstract (vs. concrete) mindset rely more on implicit associations in gist memory. Experiment 4a tests hypothesis 3 that the rating polarity effect is more likely to manifest for people in an abstract versus concrete mindset. Participants are induced with a concrete versus abstract mindset through a construal mindset manipulation (Freitas et al. 2004) before completing the same product evaluation task as in experiment 3.

Method

Participants. One hundred seventy-two US-based MTurk participants (33% female; average age: 31.2 years) participated in this experiment in return for $1.50. They were randomly assigned to one of four conditions: rating polarity (consistent vs. inconsistent) × construal mindset (abstract vs. concrete).

Procedure. The basic procedure was similar to the previous experiment, with the addition of a construal mindset manipulation after the pre-evaluation comprehension task and before the key dependent measures (purchase intention, post-evaluation comprehension task, rating polarity typicality, mood, involvement, demographics). Participants completed a mindset-priming manipulation similar to that developed by Freitas et al. (2004) and were told it was a thought-listing task that would help design future experiments. Those in the abstract mindset condition were asked to answer a series of increasingly high-level questions about *why* they would "improve and maintain one's general knowledge," while those in the concrete mindset condition were asked to answer a series of increasingly low-level questions about *how* they would engage in this same behavior.

TABLE 2

EXPERIMENTS 4A AND 4B REGRESSION RESULTS: MODERATION BY CONSTRUAL MINDSET

	Experiment 4a				Experiment 4b			
	β	SE	*t*	*p*	*β*	SE	*t*	*p*
Intercept	0.76	0.16	4.67	<.01	0.95	0.12	7.93	<0.01
Simple effects								
Rating polarity	0.12	0.25	0.50	0.62	1.01	0.17	5.97	<0.01
Quality level	0.91	0.04	23.11	<.01	0.84	0.03	27.93	<0.01
Mindset*	–0.10	0.24	–0.44	0.66	–0.06	0.03	–1.97	0.05
Two-way interaction effects								
Rating polarity × quality level	–0.10	0.06	–1.71	0.09	–0.28	0.04	–6.49	<0.01
Rating polarity × mindset	0.76	0.34	2.22	0.03	0.14	0.04	3.68	<0.01
Quality × mindset	0.06	0.06	0.96	0.34	0.01	0.01	1.41	0.16
Three-way interaction								
Rating polarity × quality level × mindset	–0.20	0.08	–2.45	0.01	–0.02	0.01	–2.56	0.01
PLANNED CONTRASTS								
Greater reliance on implicit memory**								
Rating polarity × quality level	–0.31	0.06	–5.27	<.01	–0.39	0.06	–6.42	<.01
Lower reliance on implicit memory***								
Rating polarity × quality level	–0.10	0.06	–1.71	0.09	–0.17	0.06	–2.73	0.01

*More abstract mindset (positive coefficient) indicates greater reliance on implicit memory.
**More abstract mindset.
***More concrete mindset.

Results

Post-Evaluation Comprehension Test. Five percent of participants in the inconsistent and 1% in the consistent rating polarity condition responded to at least one of the two rating-pole comprehension test questions incorrectly. These participants were excluded from the analysis.

Rating Polarity Typicality. The majority of the participants indicated that the bigger-is-better numerical association is more typical (94%) than the smaller-is-better one (3%), while 3% were not sure, confirming that the bigger-is-better association is dominant in implicit memory.

Purchase Intentions. We analyzed the data using a regression model, dummy-coding rating polarity (consistent = 0, inconsistent = 1) and construal mindset (concrete = 0, abstract = 1), and coding quality level as a continuous variable. Purchase intentions were submitted to a repeated-measures regression analysis using PROC MIXED in SAS with the three independent variables (rating polarity, mindset, and quality level) and their two- and three-way interaction terms as predicting variables (see table 2). The simple effect of quality level was significant ($\beta = .91$, $p < .01$), as was the two-way rating polarity × mindset interaction ($\beta = .76$, $p = .03$); the two-way rating polarity × quality level interaction was marginally significant ($\beta = -.10$, $p = .09$). Most importantly, the three-way rating polarity, mindset, and quality level interaction was significant ($\beta = -.20$, $p = .01$). Other effects were not significant (*p*s $> .34$; see table 2).

To examine the significant three-way interaction, we probed the two-way quality level × rating polarity interaction for abstract versus concrete mindsets. We find that the coefficient for the quality level × rating polarity interaction is statistically significant for those in an abstract mindset who rely more on implicit memory ($\beta = -.31$, $p < .01$). However, the magnitude of the coefficient was smaller and marginally significant for participants in a concrete mindset, who were less likely to rely on implicit memory ($\beta = -.10$, $p = .09$; see table 2). These results suggest that the rating polarity effect is indeed less likely to manifest for those participants in a more concrete (vs. more abstract) mindset who rely less on implicit memory, supporting hypothesis 3.

EXPERIMENT 4B: MODERATION BY CHRONIC CONSTRUAL MINDSET

It is possible that the mindset manipulation, which was administered before the main experiment, might have introduced an inadvertent confound in the previous study. To rule out any such possibility and provide further evidence that reliance on implicit versus explicit memory moderates the rating polarity effect, in this experiment we measured construal mindset using an established scale (Vallacher and Wegner 1989), rather than manipulating it.

Method

Participants. Two hundred nine undergraduates at a college in the northeast United States participated in an online study in exchange for $5 (43% female; average age: 20.1 years).

Procedure. The procedure for this experiment was identical to the previous experiment, except that we measured mindset after the product evaluation task using the Vallacher and Wegner (1989) Behavioral Identification Form (BIF) and before the post-evaluation comprehension task and measures of mood, attention, and demographics. Participants were presented with a series of 24 different behaviors (e.g., making a list), each with a more abstract response (getting organized) and a more concrete response (writing things down), and asked to indicate which of the two descriptions of the behavior they preferred.

Results

Post-Evaluation Comprehension Test. Three percent of participants in the inconsistent and 1% in the consistent rating polarity condition responded to at least one rating-pole comprehension test question incorrectly. These participants were excluded from the analysis.

Rating Polarity Typicality. The vast majority of participants indicated that the bigger-is-better numerical association is more typical (96%) than the smaller-is-better association (1%), and 3% of participants indicated they were not sure. To confirm that chronic mindset did not influence which rating polarity was seen as more typical, we conducted a logistic regression to confirm that there was no significant effect of chronic mindset, rating polarity, or the interaction between the two on perceived typicality ($ps > .30$).

Effect of Rating Polarity on Mindset. Participants' responses to the BIF study were coded as concrete = 0 and abstract = 1, and summed across the 24 responses, resulting in a continuous BIF score ranging from 0 to 24 ($M = 12.12$, SD = 4.57). This variable was mean-centered for the key analyses described below. A one-way ANOVA with rating polarity as the independent measure and construal mindset as the dependent measure confirmed that rating polarity did not have a significant effect on mindset ($p > .23$).

Purchase Intention. We again analyzed the data using a regression model, dummy-coding rating polarity (consistent as 0, inconsistent as 1), entering chronic mindset as a continuous variable (mean-centered BIF score), and coding quality level as a continuous variable. Purchase intentions were submitted to a repeated-measures regression analysis with the three independent variables (rating polarity, quality level, chronic mindset) and their two- and three-way interaction terms as predicting variables using PROC MIXED in SAS (see table 2).

The simple effects of quality level ($\beta = .84$, $p < .01$), rating polarity ($\beta = 1.01$, $p < .01$), and mindset ($\beta = -.06$, $p = .05$) were significant. The rating polarity × quality level ($\beta = -.28$, $p < .01$) and rating polarity × mindset ($\beta = .14$, $p < .01$) two-way interactions were significant. The two-way quality level × mindset interaction was not significant ($p > .15$). Most importantly, the three-way interaction was significant ($\beta = -.024$, $p = .01$). Follow-up spotlight tests (Aiken and West 1991) at one standard deviation above and below the mean of chronic mindset revealed that for those with a more chronically abstract mindset (more likely to rely on implicit memory), the two-way interaction between rating polarity and quality level was significant ($\beta = -.39$, $p < .001$; see table 2). For those with a more chronically concrete mindset (less likely to rely on implicit memory), the two-way rating polarity × quality level interaction was also significant ($\beta = -.17$, $p < .01$), but the magnitude of the coefficient was lower, indicating an attenuation of the rating polarity effect, consistent with hypothesis 3. Follow-up floodlight analyses (Spiller et al. 2013) are included in web appendix G and provide further evidence that the coefficient representing the rating polarity effect and its significance decrease for participants with an increasingly chronic concrete mindset.

EXPERIMENT 5: MODERATION BY INDIVIDUAL RELIANCE ON IMPLICIT MEMORY

The previous two experiments indirectly manipulate or measure reliance on implicit versus explicit memory. In experiment 5, we directly examine the role of reliance on implicit versus explicit memory in driving the rating polarity effect. Specifically, we test that the rating polarity effect will be stronger for people who rely to a greater extent on implicit memory (hypothesis 4). To test this, we use the well-established dual-task paradigm for measuring individual reliance on different types of memory (De Neys 2006; Jacoby 1991). In this paradigm, participants are concurrently given two tasks where one task requires implicit memory and another task requires explicit memory—performance on each of the two *simultaneous* tasks determines individuals' reliance on implicit and explicit memory (De Neys 2006; Jacoby 1991). We created a computerized dual task, administered online and modeled after the prevailing dual tasks, in which the computer coded the responses (see web appendix B). In this dual task, participants had to simultaneously complete a true/false task (which measured reliance on implicit memory) and a recall task (which measured reliance on explicit memory).

More specifically, in the dual task, participants were shown a series of simple arithmetic facts and had to submit a true or false judgment (e.g., "7 – 5 = 1"). Participants had to make the judgment in 3 seconds, while concurrently engaging in a recall task. For the recall task, participants had to memorize the series of digits on the right-hand side of the equations (e.g., 1 in the equation above) and report them at the end of a series of true or false questions. Thus, when making the true/false judgments in this dual-task paradigm, participants had neither time nor working memory resources, and had to rely on pattern recognition stored in their implicit memory. Prior research has shown that when you look at "7 – 5 = 1" you can intuitively sense, using implicit memory, whether this statement is true or false, without actually subtracting (Schunn et al. 1997). Success in the true/false task, measured as the number of correct responses to the true/false questions in this dual-task paradigm, is a direct measure of participants' reliance on implicit memory. Success in the recall task, measured as the number of correct responses to the recall questions in this dual-task paradigm, is a measure of participants' reliance on explicit memory. We predicted that the rating polarity effect would be stronger for those participants who receive a higher score in the true/false task (indicating that they rely more on implicit memory). Because we posit that the rating polarity effect stems from spontaneous interference from implicit memory, we predict that subjects' performance on the implicit memory task, but not the explicit memory task, will moderate the rating polarity effect.

Method

Participants. Two hundred two US-based MTurk participants (38% female; average age: 33.4 years) participated in this experiment in return for $1.50.

Procedure. This experiment was identical to experiment 4b, except that the task measuring implicit and explicit memory replaced the BIF measurement. After the main experiment, participants responded to an ostensibly unrelated dual-task test where they were told that they would complete two tasks simultaneously—a true/false task and a recall task. They were shown a series of simple arithmetic facts. For the true/false task, they had to make a determination of whether the arithmetic fact was true or false within 3 seconds. For example, in a block with four facts, participants saw the following four arithmetic facts, one at a time, and they had to indicate whether each of them was true or false: "7 – 5 = 1"; "2 + 3 = 5"; "2 + 5 = 6"; "4 + 5 = 8" (see web appendix B for stimuli examples). In this block the correct responses to the four equations are false, true, false, and false, respectively. While making these judgments, participants had to simultaneously complete a memorization task. For the memorization task, they needed to memorize the number on the right-hand side of the equation in the order they saw it and then recall these numbers in order. For example, while completing the true/false task for the first block, they had to memorize the right-side digits of the equations listed above and recall the string of digits 1568 at the end of the block. Participants completed eight such blocks of equations with increasing spans of the recall task; there were four, five, six, six, seven, seven, eight, and eight equations in each succeeding block (see web appendix B). Thus, over eight blocks, participants made a total of 51 true/false judgments and simultaneously responded to 51 explicit recall questions. Before completing the main dual task of these eight blocks, participants received three practice blocks, with feedback, to understand the task.

Results

Post-Evaluation Comprehension Test. Two percent of participants in the inconsistent and 2% in the consistent rating polarity condition responded to at least one rating-pole comprehension test question incorrectly. As in the previous studies, these participants were excluded from the analysis. Additionally, one participant who got all 51 true/false judgments incorrect was excluded from the analyses.

Rating Polarity Typicality. The vast majority of the participants indicated that the bigger-is-better numerical association is more typical (84%) than the smaller-is-better one (16%), confirming that the bigger-is-better association is more dominant in implicit memory.

Effect of Rating Polarity on Implicit Memory Task Performance. Participants' propensity to rely on implicit memory was measured by the number of correct responses to the true/false task. Larger numbers indicate a greater propensity to rely on implicit memory. These scores ranged from 18 to 51 with a mean of 42.65 and standard deviation of 6.80. A one-way ANOVA with rating polarity as the independent variable and performance on the true/false task as the dependent variable confirmed that rating polarity did not have a significant effect on performance on the true/false task ($F < 1$).

Purchase Intentions. We analyzed purchase intentions using a regression model, dummy-coding rating polarity (consistent = 0, inconsistent = 1), entering performance on the implicit memory task as a continuous variable (mean-centered), and coding quality level as a continuous variable (inadequate = 1, adequate = 2, fair = 3, good = 4, very good = 5). Purchase intentions were submitted to a repeated-measures regression analysis with the three independent variables (rating polarity, quality level, implicit memory task performance) and their two- and three-way interaction terms as predicting variables, using PROC MIXED in SAS (see table 3).

The simple effects of rating polarity ($\beta = .60$, $p < .01$), quality level ($\beta = 1.06$, $p < .01$), and propensity to rely on implicit memory ($\beta = -.05$, $p < .01$) were significant. The two-way interactions of rating polarity × quality level ($\beta = -.14$, $p < .01$) and quality level × implicit memory reliance ($\beta = .010$, $p < .01$) were significant. The two-way interaction between rating polarity × implicit memory

TABLE 3

EXPERIMENT 5 REGRESSION RESULTS: MODERATION BY RELIANCE ON IMPLICIT MEMORY

	β	SE	*t*	*p*
Intercept	0.139	0.115	1.21	0.23
Simple effects				
Rating polarity	0.596	0.161	3.7	<.01
Quality level	1.057	0.026	40.05	<.01
Implicit memory reliance	–0.047	0.017	–2.85	<.01
Two-way interaction effects				
Rating polarity × quality level	–0.144	0.037	–3.89	<.01
Rating polarity × implicit memory reliance	0.024	0.024	1.01	0.31
Quality × implicit memory reliance	0.01	0.004	2.73	0.01
Three-way interaction				
Rating polarity × quality level × implicit memory reliance	–0.011	0.005	–2.00	0.05
PLANNED CONTRASTS				
Greater propensity to use implicit memory* Rating Polarity × quality level	–0.218	0.052	–4.17	<.01
Lower propensity to use implicit memory** Rating polarity × quality level	–0.07	0.052	–1.33	0.18

**One SD above the mean for each memory type.
***One SD below the mean for each memory type.

reliance was not significant. Most importantly, the three-way interaction was significant ($\beta = -.01$, $p = .05$).

Follow-up spotlight tests (Aiken and West 1991) at one standard deviation above and below average performance scores on the implicit memory task revealed that for those who performed above average on the implicit memory task, the two-way rating polarity × quality level interaction representing the rating polarity effect was significant ($\beta = -.22$, $p < .001$; see table 3). However, for those participants who performed below average on the implicit memory task, the two-way rating polarity × quality level interaction was not significant ($\beta = -.07$, $p = .18$), indicating an attenuation of the rating polarity effect. These results support hypothesis 4. We performed a similar regression analysis using performance on the explicit memory performance task rather than the implicit memory task as the dependent measure. The three-way interaction between rating polarity, quality level, and explicit memory reliance was not statistically significant ($\beta = .002$, $p = .59$), indicating that although a direct measure of reliance on implicit memory moderates the rating polarity effect, the direct measure of reliance on explicit memory does not.

DISCUSSION OF EXPERIMENTS 4A, 4B, AND 5

Together, experiments 4a, 4b, and 5 provide evidence that the rating polarity effect is more likely to occur when people rely more on their implicit memory. Experiments 4a and 4b show that participants under an abstract (vs. concrete) construal mindset, which increases reliance on implicit memory, are more likely to exhibit the rating polarity effect. The size of the coefficient for the rating polarity × quality level interaction, which represents the extent of interference from the implicit numerical association when using a rating format with inconsistent rating polarity, is smaller for those participants who are either induced with a concrete mindset (experiment 4a) or chronically adopt a more concrete mindset (experiment 4b).

In experiment 5, we directly measured the use of implicit versus explicit memory using the dual-task paradigm and found that the rating polarity effect was moderated by the propensity to use implicit memory: the rating polarity effect was stronger for those participants who relied on implicit memory to a greater extent during the dual task. Furthermore, reliance on explicit memory does not moderate the rating polarity effect, providing further evidence that it is an effect driven by spontaneous interference from the numerical association in implicit memory. The strongest moderating effect comes from directly measuring individuals' use of implicit memory (experiment 5). Mood and involvement cannot account for these effects (web appendix D). These experiments support hypotheses 3 and 4, demonstrating that people are more likely to experience spontaneous proactive interference from an implicit numerical association in memory when in a mindset that encourages reliance on implicit memory or if their individual chronic propensity is to rely more on implicit memory.

GENERAL DISCUSSION

Across seven experiments, we provide evidence of a new form of proactive interference between numerical associations: the rating polarity effect. The first three experiments demonstrated the effect, supporting hypothesis 1: evaluating a product using a rating format with a numerical association opposite the one that people hold in implicit memory results in evaluations that are less responsive to differences in quality level. This effect is robust to consumer evaluations ranging across bidding behavior (paradigm pretest), visual perception (experiment 1), purchase intent (experiments 2a, 3, 4a, 4b, and 5), and willingness to pay (experiment 2b). It manifests across a variety of response alternatives (open-ended, horizontal and vertical orientation, with and without numerical anchors, binary and scale). Confusion, forgetting, mood, and involvement are ruled out as potential alternative explanations. The effect occurs for single and repeated evaluations and is not moderated by order or need for cognition, demonstrating the spontaneous and unintended effect of the implicit numerical association in memory.

To understand the mechanism behind the rating polarity effect and why this proactive interference effect persists even when participants are well aware of the rating polarity employed, we examine the role of implicit memory. If the rating polarity effect is caused by the spontaneous use of the implicit numerical association in memory, then the rating polarity effect should be moderated by factors that moderate reliance on implicit memory. Supporting hypothesis 2, experiment 3 demonstrates that for participants from a country with an opposite numerical association as a cultural norm, the effect of bigger-is-better versus smaller-is-better rating polarity reverses. In addition, for those participants with a more flexible set of implicit numerical associations in memory, the rating polarity effect does not manifest, as they do not experience proactive interference from a strong implicit numerical association. Further illustrating the role of implicit numerical associations behind the rating polarity effect, this proactive interference of numerical associations diminishes when people rely less on implicit memory, whether it is based on a mindset that leads people to rely less on implicit memory (experiments 4a and 4b, supporting hypothesis 3) or on individual differences in relying on implicit memory as directly measured through a dual-task paradigm (experiment 5, supporting hypothesis 4).

Theoretical and Managerial Contributions

The current research presents new insights into numerical cognition and makes three theoretical contributions. First, we characterize a new type of implicit numerical association that differs depending on cultural context—the association between the magnitude of a number and an evaluative judgment. Culture can influence whether consumers have a bigger-is-better versus smaller-is-better numerical association in implicit memory, and this implicit association can spontaneously influence the judgments they make when using a rating polarity with an opposite numerical association. Second, our research demonstrates the persistence of interference effects that stem from an implicit association in memory that has been culturally cultivated over time. The rating polarity effect manifests over different types of evaluations, over multiple consumer evaluations, and even when measures of difficulty and need for cognition (NFC) are taken into account.

Third, our research also contributes to the memory literature by identifying novel moderators of proactive interference, based on unintentional reliance on implicit memory. We show that the strength of the implicit numerical association in memory moderates the effect of proactive interference. Furthermore, construal mindset (both manipulated and measured) can moderate proactive interference, and add to the burgeoning literature on how cognitive mindsets change memory processes. In particular, in contrast to work that demonstrates that abstract construal mindsets can attenuate interference when explicit memory interferes with implicit memory (Kyung and Thomas 2016), this research demonstrates that concrete construal mindsets can attenuate interference when implicit memory interferes with explicit memory. Thus, whether a concrete versus abstract construal mindset can attenuate memory interferences depends on which form of memory is causing the interference effect. In addition, we show that individual propensity to rely on implicit memory (as measured by performance on the implicit memory portion of the dual-task exercise) can moderate proactive interference. If one thinks of the implicit numerical association (bigger is better) as a culturally induced metaphor (Casasanto 2014), then our research introduces the notion of interference that stems from metaphorical incongruity. Future research could examine other metaphors for interference.

From a managerial perspective, our research suggests that rating polarity is an important element of market research design in cross-cultural contexts and for the presentation of results. Presenting information using rating polarity inconsistent with the numerical association in implicit memory can make consumers less responsive to differences in numeric quality ratings of products. Therefore, retailers should translate numeric ratings into a format that is consistent with the numerical association that their target customers have in memory. Similarly, any decision maker that encounters numeric information utilizing rating polarity opposite the numerical association they hold in implicit memory should be aware of the potential impact on their judgments—for example, admissions officers or human resource managers who use grade point average information across cultures; policy makers making global decisions based upon cross-cultural research, such as life satisfaction surveys where some countries use smaller-is-better rating polarity and some do not (Diener, Kahneman, and Helliwell 2010); and reviewers evaluating the $30 billion in research grants for the National Institutes of Health, where proposals are evaluated in a format in which 1 = exceptional and 9 = poor.

Finally, our work underscores the importance of understanding culturally implicit associations that might bias people's judgments and how difficult it can be to overcome these implicit associations, such as the implicit association between "bigger" and "better" in the United States. Thus, we encourage managers and researchers to understand what these implicit associations might be and how they can influence judgments.[1]

DATA COLLECTION INFORMATION

All study designs and analyses were discussed and agreed upon by all three authors. The first author supervised the data collection for experiments 1a and 3b through Dartmouth College students, experiment 1b through Cornell University students, study 2 through a Qualtrics panel of German participants, and studies 2a, 2b, 3a, and 4 through the MTurk online panel. The first and second authors separately and jointly analyzed the data from these studies, confirming all results.

REFERENCES

Adaval, Rashmi and Kent B. Monroe (2002), "Automatic Construction and Use of Contextual Information for Product and Price Evaluations," *Journal of Consumer Research*, 28 (4), 572–88.

Aiken, Leona S. and Stephen G. West (1991), *Multiple Regression: Testing and Interpreting Interactions*, London: Sage Publications.

Alter, Adam L. and Daniel M. Oppenheimer (2009), "Uniting the Tribes of Fluency to Form a Metacognitive Nation," *Personality and Social Psychology Review*, 13 (3), 219–35.

Bagchi, Rajesh and Derick F. Davis (2012), "$29 for 70 items or 70 items for $29? How Presentation Order Affects Package Perceptions," *Journal of Consumer Research*, 39 (1), 62–73.

Blankenship, Albert B. and Paul L. Whitely, (1941), "Proactive Inhibition in the Recall of Advertising Material," *Journal of Social Psychology*, 13 (2), 311–22.

Burke, Raymond R. and Thomas K. Srull (1988), "Competitive Interference and Consumer Memory for Advertising," *Journal of Consumer Research*, 15 (6), 55–68.

Cacioppo, John T., Richard E. Petty, and Chuan Feng Kao (1984), "The Efficient Assessment of Need for Cognition," *Journal of Personality Assessment*, 48 (3), 306–7.

Casasanto, Daniel (2014), "Experiential Origins of Mental Metaphors: Language, Culture, and the Body," in *The Power of Metaphor: Examining Its Influence on Social Life*, ed. Mark J. Landau, Michael D. Robinson, and Brian P. Meier, Washington, DC: American Psychological Association.

Dehaene, Stanislas, Serge Bossini, and Pascal Giraux (1993), "The Mental Representation of Parity and Number Magnitude," *Journal of Experimental Psychology: General*, 12 (3), 371–96.

[1] We conclude this paper as we have found that the association between "bigger" and better" does not necessarily apply to the length of academic papers.

De Neys, Wim (2006), "Dual Processing in Reasoning Two Systems but One Reasoner," *Psychological Science*, 17 (5), 428–33.

Diener, Ed, Daniel Kahneman, and John Helliwell (2010), *International Differences in Well-Being*, New York: Oxford University Press.

Florack, Arnd, Martin Scarabis, and Herbert Bless (2001), "When Do Associations Matter? The Use of Automatic Associations toward Ethnic Groups in Person Judgments," *Journal of Experimental Social Psychology*, 37 (6), 518–24.

Freitas, Antonio L., Peter Gollwitzer, and Yaacov Trope (2004), "The Influence of Abstract and Concrete Mindsets on Anticipating and Guiding Others' Self-Regulatory Efforts," *Journal of Experimental Social Psychology*, 40 (6), 739–52.

Fukukura, Jun, Melissa J. Ferguson, and Kentaro Fujita (2013), "Psychological Distance Can Improve Decision Making under Information Overload via Gist Memory," *Journal of Experimental Psychology: General*, 142 (3), 658–65.

Gilbert, Daniel T. and J. Gregory Hixon (1991), "The Trouble of Thinking: Activation and Application of Stereotypic Beliefs," *Journal of Personality and Social Psychology*, 60 (4), 509–17.

Graf, Peter and Daniel L. Schacter (1987), "Selective Effects of Interference on Implicit and Explicit Memory for New Associations," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13 (1), 45–53.

Jacoby, Larry L. (1991), "A Process Dissociation Framework: Separating Automatic From Intentional Uses of Memory," *Journal of Memory and Language*, 30 (5), 513–41.

Jahoda, Gustav (1969), *The Psychology of Superstition*, Oxford, UK: Penguin Books.

Jonides, John and Derek E. Nee (2006), "Brain Mechanisms of Proactive Interference in Working Memory," *Neuroscience*, 139 (1), 181–93.

Keller, Kevin Lane (1987), "Memory Factors in Advertising: The Effect of Advertising Retrieval Cues on Brand Evaluations," *Journal of Consumer Research*, 14 (3), 316–33.

—— (1991), "Memory and Evaluations in Competitive Advertising Environments," *Journal of Consumer Research*, 17 (3), 463–76.

Keller, Kevin Lane, Susan E. Heckler, and Michael J. Houston (1998), "The Effects of Brand Name Suggestiveness on Advertising Recall," *Journal of Marketing* 62 (1), 48–57.

Keppel, Geoffrey and Benton J. Underwood (1962), "Proactive Inhibition in Short-Term Retention of Single Items," *Journal of Verbal Learning and Verbal Behavior*, 1 (3), 153–61.

King, Dan and Chris Janiszewski (2011), "The Sources and Consequences of the Fluent Processing of Numbers," *Journal of Marketing Research*, 48(2), 327-41.

Kyung, Ellie J. and Manoj Thomas (2016), "When Remembering Disrupting Knowing: Blocking Implicit Price Memory," *Journal of Marketing Research*, 53 (6), 937–53.

Lee, Angela Y. and Aparna A. Labroo (2004), "The Effect of Conceptual and Perceptual Fluency on Brand Evaluation," *Journal of Marketing Research*, 41 (2), 151–65.

McCrea, Sean M., Frank Wieber, and Andrea L. Myers (2012), "Construal Level Mind-Sets Moderate Self-and Social Stereotyping," *Journal of Personality and Social Psychology*, 102 (1), 51–68.

Mishra, Himanshu, Arul Mishra, and Dhananjay Nayakankuppam (2006), "Money: A Bias for the Whole," *Journal of Consumer Research*, 32 (4), 541–49.

Monga, Ashwani and Rajesh Bagchi (2012), "Years, Months, and Days vs. 1, 12, and 365: The Influence of Units Versus Numbers," *Journal of Consumer Research*, 39 (1), 185–98.

Monroe, Kent B. and Angela Y. Lee (1999), "Remembering versus Knowing: Issues in Buyers' Processing of Price Information," *Journal of the Academy of Marketing Science*, 27 (2), 207–25.

Mussweiler, Thomas and Birte Englich (2005), "Subliminal Anchoring: Judgmental Consequences and Underlying Mechanisms," *Organizational Behavior and Human Decision Processes*, 98 (2), 133–43.

Raghubir, Priya and Joydeep Srivastava (2009), "The Denomination Effect," *Journal of Consumer Research*, 36 (4), 701–13.

Rim, SoYon, James S. Uleman, and Yaacov Trope (2009), "Spontaneous Trait Inference and Construal Level Theory: Psychological Distance Increases Nonconscious Trait Thinking," *Journal of Experimental Social Psychology*, 45 (5), 1088–97.

Roediger, Henry L. (1990), "Implicit Memory: Retention without Remembering," *American Psychologist*, 45 (9), 1043–56.

Schacter, Daniel L. (1987), "Implicit Memory: History and Current Status," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13 (3), 501–18.

Schunn, Christian D., Lynne M. Reder, Adisack Nhouyvanisvong, Daniel Richards, and Philip J. Stroffolino (1997),"To Calculate or Not to Calculate: A Source Activation Confusion Model of Problem Familiarity's Role in Strategy Selection," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23 (1), 3–29.

Schwarz, Norbert (2004), "Meta-Cognitive Experiences in Consumer Judgment and Decision Making," *Journal of Consumer Psychology*, 14 (4), 332–48.

Smith, Pamela K. and Yaacov Trope (2006), "You Focus on the Forest When You're in Charge of the Trees: Power Priming and Abstract Information Processing," *Journal of Personality and Social Psychology*, 90 (4), 578–96.

Spiller, Stephen A., Gavan J. Fitzsimons, John G. Lynch, and Gary H. McClelland (2013), "Spotlights, Floodlights, and the Magic Number Zero: Simple Effects Tests in Moderated Regression," *Journal of Marketing Research*, 50 (2), 277–88.

Thomas, Manoj and Vicki G. Morwitz (2009), "Heuristics in Numerical Cognition: Implications for Pricing," in *Handbook of Pricing Research in Marketing*, ed. Vithala R. Rao, Cheltenham, UK and Northampton, MA: Edward Elgar.[*WorldCat*]

Trope, Yaacov and Nira Liberman (2003), "Temporal Construal," *Psychological Review*, 110 (3), 403–21.

Tversky, Amos and Daniel Kahneman (1974), "Judgment under Uncertainty: Heuristics and Biases," *Science*, 185 (4157), 1124–31.

Unnava, H. Rao and Deepak Sirdeshmukh (1994), "Reducing Competitive Ad Interference," *Journal of Marketing Research*, 31(3), 403–11.

Vallacher, Robin R. and Daniel M. Wegner (1989), "Levels of Personal Agency: Individual Variation in Action Identification," *Journal of Personality and Social Psychology*, 57 (4), 660–71.

Wickens, Delos D., David G. Born, and Charles K. Allen (1963), "Proactive Inhibition and Item Similarity in Short-Term Memory," *Journal of Verbal Learning and Verbal Behavior*, 2 (5), 440–45.

Yates, J. K. (2007), *Global Engineering and Construction*, Hoboken, NJ: John Wiley & Sons.