

When Bigger is Better (and When it is Not): Implicit Bias in Numeric Judgments

ELLIE J. KYUNG

MANOJ THOMAS

ARADHNA KRISHNA*

Web Appendix Contents

- Web Appendix A: Pretest—Bidding Behavior in an Auction
- Web Appendix B: Experiment Stimuli
- Web Appendix C: Experiment Results (means, standard deviation, and confidence intervals; graphs of means and confidence intervals)
- Web Appendix D: Ruling out possible Alternative Explanation (means and standard deviation; analyses details)
- Web Appendix E: Analyses with Reaction Time as Covariate
- Web Appendix F: Order Effect Analyses
- Web Appendix G: Experiment 4B Floodlight Analysis

WEB APPENDIX A:

PRETEST—BIDDING BEHAVIOR IN AN AUCTION

In this pretest experiment, we test the rating polarity effect (testing H1) in a sealed-bid auction that entailed financial consequences. We examined whether rating polarity influenced participants' bids. Comprehension, involvement, mood, and need-for-cognition were also measured to rule out potential alternative explanations.

Method

Participants. Seventy-two undergraduates from a U.S. university (60% female; average age: 20.5 years) participated in the computer study in exchange for \$4. Participants were randomly assigned to either the consistent (bigger-is-better) or inconsistent (smaller-is-better) rating polarity condition.

Procedure. Participants were told that they would be offered the opportunity to bid for a Japanese insulated travel mug that was not available for purchase locally, and that the participant who placed the winning bid would receive the travel mug. They could use all or some portion of their \$4 payment to bid on the mug. If the winning bid was less than \$4, the winner would receive the difference between the payment and their bid price. They were also told they could bid more than \$4, but that they would personally have to pay the difference between their bid price and the \$4 payment to receive the mug. Participants were asked to confirm [*Yes / No*] that they understood the instructions for the bidding process before proceeding.


Before they were shown the product information, participants were told that a quality rating for the mug would be provided by a reputable consumer welfare agency outside of the U.S.—an agency known for its evaluation of consumer products. They were asked to carefully read the meaning of the ratings provided. 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 the rating had an opposite polarity of 1 = *very good* and 7 = *unsatisfactory*. They were asked to confirm that they understood this rating format [*Yes / No*].

PRETEST EXPERIMENT STIMULI

Product features:

- 16 oz.
- Stainless steel interior with vacuum insulation
- Stainless steel exterior
- Keeps beverages hot or cold for 6 hours
- Compact design
- Flip lid with safety lock to prevent spills
- BPA-free
- Easy clean
- Produced by reputable Japanese company

Quality Rating: 6.1



Continue

Pre-evaluation Comprehension Test. Participants were then asked the meaning of “1” and “7” rating poles for the quality ratings [*very good, unsatisfactory*] before proceeding. If they responded incorrectly, a message came up on the screen asking them to correct their response. Participants could proceed only after they had answered these questions correctly. This was done to ensure that the results did not stem from inattention or miscomprehension of the rating poles.

Bid. Participants were then shown a photograph and descriptive information about the insulated mug, which included a quality rating (see Web Appendix B). The quality rating was 6.1 in the consistent rating polarity condition (bigger-is-better) and 1.9 in the inconsistent rating polarity condition (smaller-is-better). Note that the two ratings are normatively identical. Participants were then asked, “How much would you like to bid on this mug?” and were required to make a bid, but could make any bid they wished, including \$0.

Post-Evaluation Comprehension Test. After entering their bid, participants were asked what the exterior of the mug was made of [*stainless steel or plastic*], the country of origin of the product [*Germany or Japan*], and 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 during the bidding process.

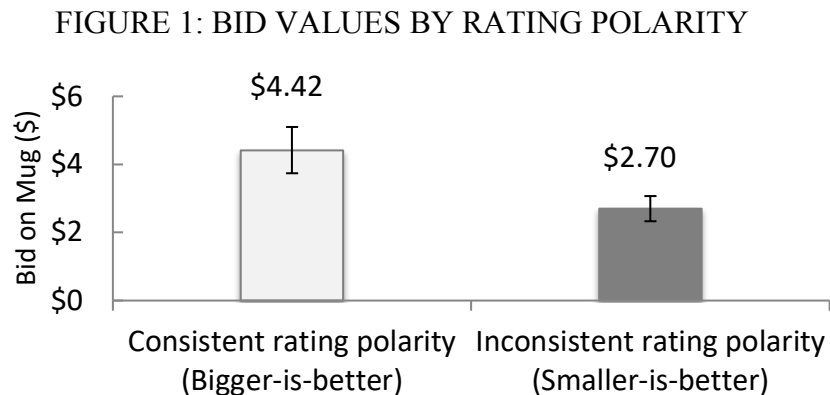
Rating Polarity Typicality. Participants were also asked to indicate which of the two numerical associations is more typical: “Higher Numbers Indicate Better Quality” or “Lower Numbers Indicate Better Quality.” They could also choose a third option labeled “not sure.”

Additional Measures. To rule out other 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), asked how involved / thoughtful / attentive they were during the study [slider scales anchored at *not at all / very*], and asked questions about their demographics.

Results

Post-Evaluation Comprehension Test. All participants answered both of the questions on the meaning of the rating poles correctly. Thus, they were neither confused about, nor did they forget, the meaning of the poles.

Monetary Bid. A one-way ANOVA revealed that, as predicted, participants in the consistent rating polarity condition bid more on the insulated travel mug ($M = \$4.42$) than those in the inconsistent rating polarity condition ($M = \$2.70$, $F(1, 70) = 4.80$, $p = .03$, see figure 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.

Rating Polarity Typicality. The majority of participants indicated that the bigger-is-better numerical association is more typical (90.3%) than the smaller-is-better one (9.7%), and thus the dominant numerical association in implicit memory.

Ruling Out Alternative Explanations. Mood: To rule out the possibility that an inconsistent rating polarity (smaller-is-better) might have put participants in a more negative mood and influenced their information processing, we averaged three mood measures into an index ($\alpha = .92$) and conducted a one-way ANOVA, which revealed no difference in mood by condition ($p = .42$). Involvement: Similarly, to rule out the possibility that participants might have been more involved when attempting to use a rating format with inconsistent rating polarity, we averaged the three measures of involvement into an index ($\alpha = .87$), and a one-way ANOVA revealed no difference in involvement by condition ($p = .60$). Need-for-cognition: We propose that the rating polarity effect manifests because participants are not aware of, and therefore cannot control for, its effects. Previous research has shown that individuals who are higher NFC are more systematic in their judgments (Florack and Scarabis 2001; Conner et al. 2007) because they are more thoughtful and reflective (Ajezen and Fishbein 2005)—for example, they are less likely to make judgments in line with negative stereotypes. However, if the rating polarity effect operates outside of people’s awareness, then high NFC individuals should not be able to overcome the effect of interference from implicit numerical associations in memory.

We conducted a linear regression 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 the bid as the dependent measure. The results revealed a significant effect of rating polarity ($\beta = -1.84, p = .03$), but the effects of NFC ($p = .29$) and the interaction between NFC and rating polarity ($p = .66$) were not significant. This suggests that the rating polarity effect manifests for people with both low and high NFC and provides some evidence that the spontaneous interference of implicit memory is non-conscious because high NFC participants cannot adjust for it.

The means and standard deviations for these additional variables (mood, involvement, NFC) across all experiments are in Web Appendix D. These measures also show that participant involvement and attention are relatively high through all of the experiments, across experimental conditions, and low involvement or attention cannot account for the results.

WEB APPENDIX B
EXPERIMENT 1 STIMULI

One kit includes:

- 20 applications (15 min each)
- 3 mouth trays
- Shade guide

Product is:

- Peroxide free
- Enamel safe

Quality Rating: 1.9



How much **WHITER** do the teeth look in the 'after' versus the 'before' photograph?

Drag the **blue knob** to the point that represents your opinion below and press "continue."

not at all whiter  **much whiter**

Continue

EXPERIMENT 2B STIMULI

Please imagine that you are shopping and that you want to buy water. You see the following product.

Böhmers Quelle

Quality Rating: 2.0

The water for your well-being



How much would you be willing to pay for this product in dollars between \$0.00 and \$6.00 ? Please choose a price from the list provided below.

- ✓ \$0.00
- \$0.50
- \$1.00
- \$1.50
- \$2.00
- \$2.50
- \$3.00
- \$3.50
- \$4.00
- \$4.50
- \$5.00
- \$5.50
- \$6.00

>>

Survey Powered By [Qualtrics](#)

EXPERIMENT 4A STIMULI

'WHY' MINDSET MANIPULATION

WHY TO IMPROVE & MAINTAIN ONE'S GENERAL KNOWLEDGE

For this thought exercise, please consider the following activity: **improving and maintaining one's general knowledge**.

Now complete the following flowchart by listing three ways in which improving and maintaining your general knowledge can help you meet your important life goals. Please start your response from the box at the bottom of the page (FIRST OUTCOME) and think of one immediate benefit of having good general knowledge. Then in the next box (SECOND OUTCOME), list the benefit of the FIRST OUTCOME and repeat the process until you complete the three outcomes.

TO BEGIN SCROLL DOWN TO THE BOTTOM OF THIS WEBPAGE AND START WITH THE FIRST OUTCOME.

(THIRD OUTCOME)

WHY?



(SECOND OUTCOME)

WHY?



START HERE (FIRST OUTCOME)

WHY?



IMPROVE AND MAINTAIN ONE'S GENERAL KNOWLEDGE

‘HOW’ MINDSET MANIPULATION

HOW TO IMPROVE & MAINTAIN ONE'S GENERAL KNOWLEDGE

For this thought exercise, please consider the following activity: how to **improve and maintain one's general knowledge**.

Now complete the following flowchart by listing three ways by which you can improve and maintain your general knowledge. Please start your response from the box at the top of the webpage and think of one immediate way or step (FIRST STEP) to improve general knowledge. Then you become more specific and think about the next step (SECOND STEP) to implement the FIRST STEP and repeat the process until you complete the three steps.

START YOUR RESPONSE WITH THE FIRST STEP.

IMPROVE AND MAINTAIN ONE'S GENERAL KNOWLEDGE

HOW?



(FIRST STEP)

HOW?



(SECOND STEP)

HOW?



(THIRD STEP)

WEB APPENDIX C: EXPERIMENT RESULTS

MEANS, STANDARD DEVIATIONS, AND +/- 95% CONFIDENCE INTERVAL

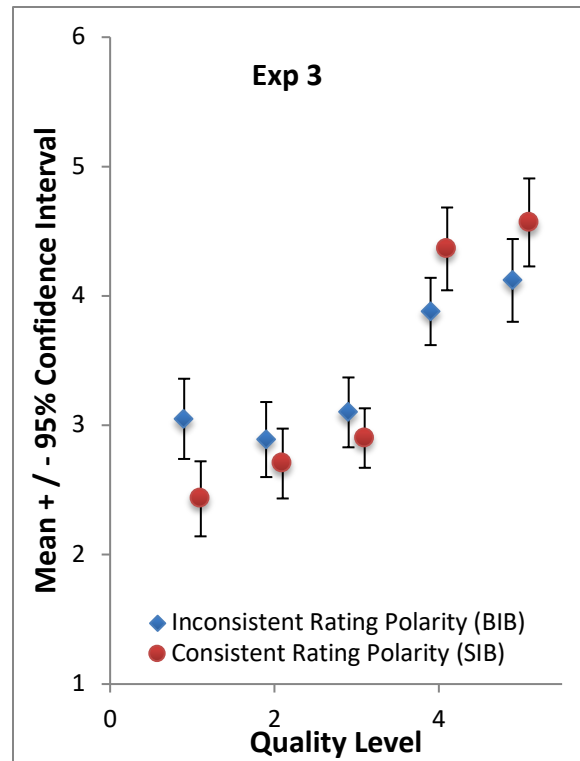
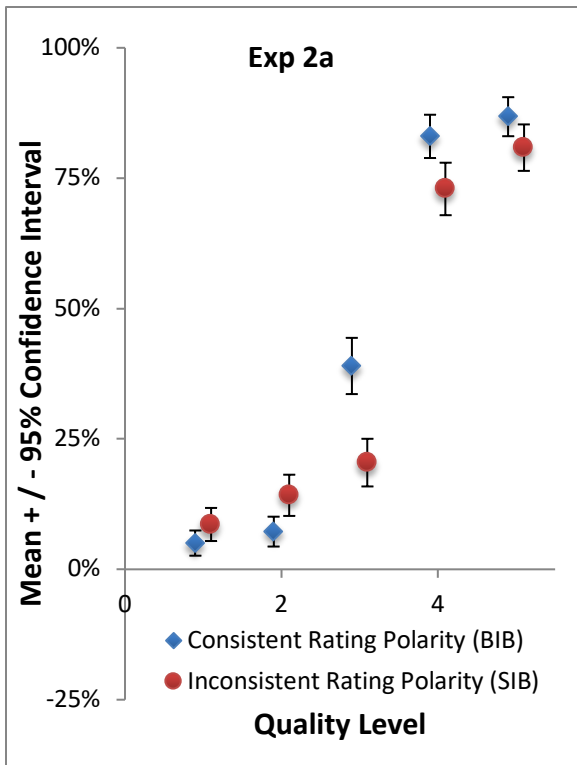
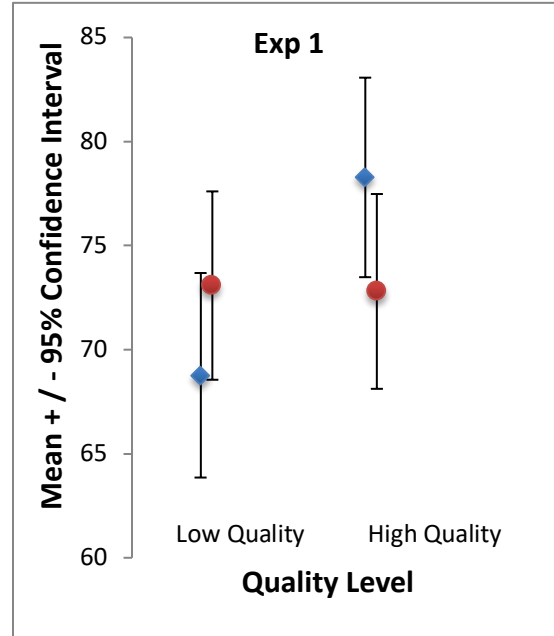
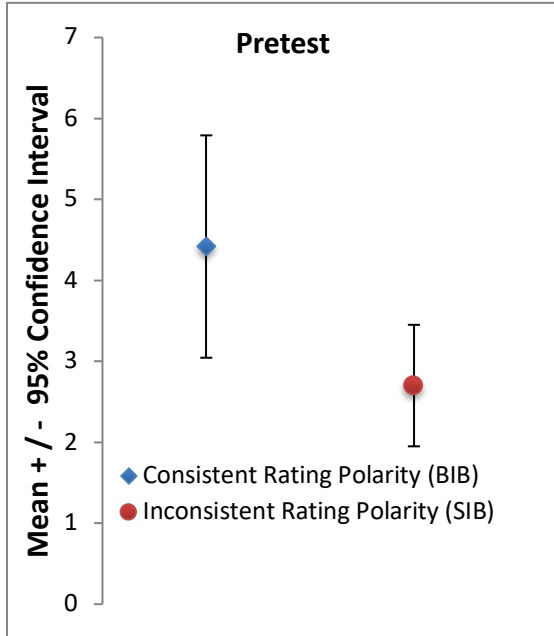
BIB					SIB				
Quality Level	Mean	SD	Confidence Interval		Quality Level	Mean	SD	Confidence Interval	
			Lower-Bound	Upper-Bound				Lower-Bound	Upper-Bound
PRETEST EXPERIMENT									
6.10	4.42	4.12	3.04	5.79	1.90	2.70	2.18	1.95	3.45
EXPERIMENT 1									
Low	68.77	16.88	63.86	73.68	1	73.08	13.92	68.55	77.61
High	78.28	13.97	73.48	83.07	2	72.80	17.27	68.12	77.48
EXPERIMENT 2A									
1	5.0%	1.2%	2.6%	7.4%	1	8.6%	1.6%	5.4%	11.8%
2	7.2%	1.5%	4.4%	10.1%	2	14.2%	2.0%	10.2%	18.1%
3	39.0%	2.7%	33.6%	44.4%	3	20.5%	2.3%	15.9%	25.0%
4	83.0%	2.1%	78.9%	87.2%	4	72.9%	2.6%	67.9%	78.0%
5	86.8%	1.9%	83.1%	90.5%	5	80.9%	2.3%	76.4%	85.3%
EXPERIMENT 2B									
1	0.73	0.82	0.59	0.88	1	1.00	1.33	0.86	1.14
2	0.97	0.79	0.83	1.12	2	1.24	1.24	1.09	1.38
3	1.72	0.94	1.57	1.86	3	1.57	0.85	1.43	1.72
4	2.33	1.08	2.19	2.47	4	2.33	0.92	2.18	2.47
5	2.75	1.34	2.61	2.90	5	2.67	1.05	2.53	2.81
EXPERIMENT 3									
1	2.43	1.71	0.15	2.14	1	3.05	1.83	0.16	2.74
2	2.70	1.60	0.14	2.43	2	2.89	1.68	0.14	2.60
3	2.90	1.31	0.11	2.68	3	3.10	1.58	0.14	2.83
4	4.36	1.83	0.16	4.05	4	3.88	1.53	0.13	3.62
5	4.57	1.97	0.17	4.23	5	4.12	1.89	0.16	3.80

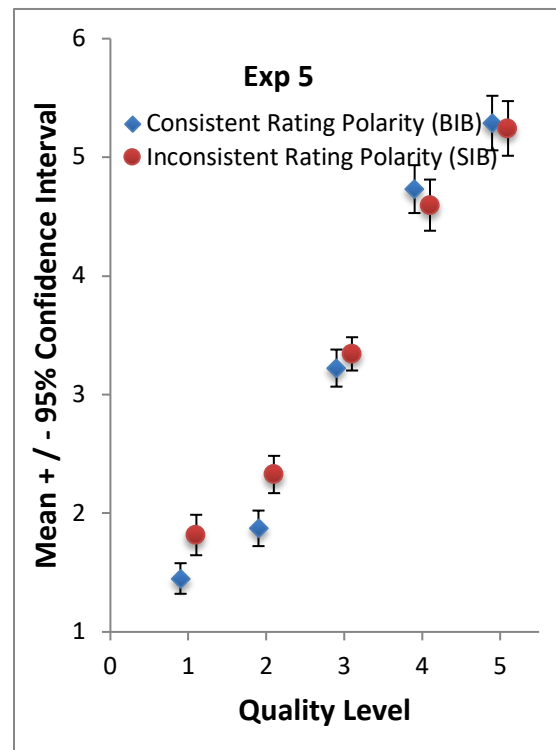
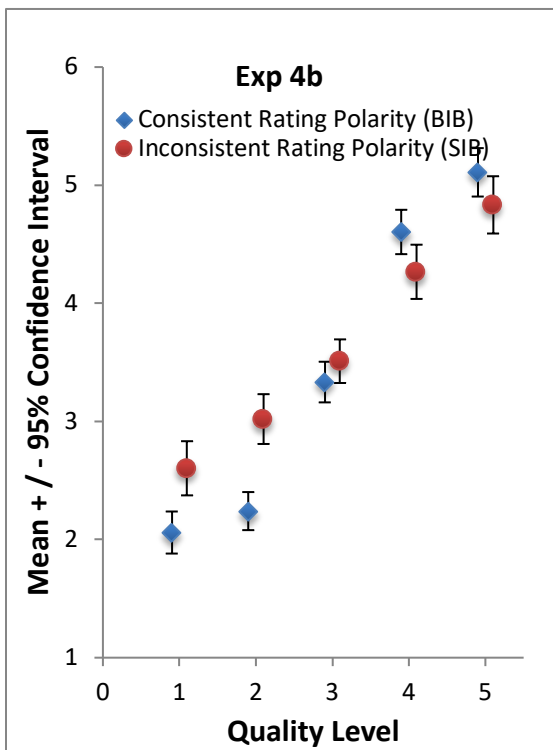
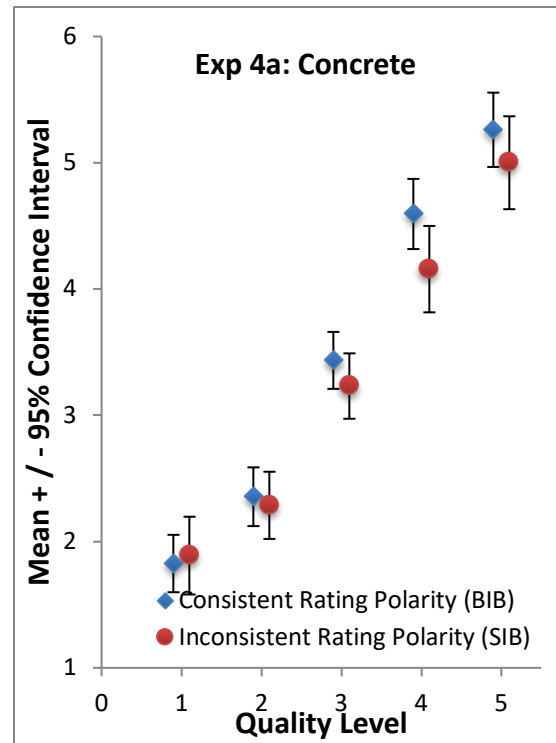
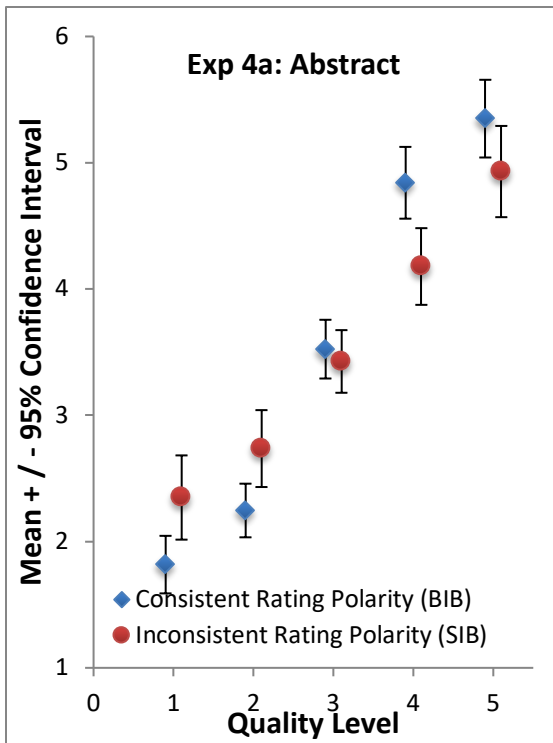
MEANS, STANDARD DEVIATIONS, AND +/- 95% CONFIDENCE INTERVAL
CONTINUED

BIB					SIB				
Quality Level	Mean	SD	Confidence Interval		Quality Level	Mean	SD	Confidence Interval	
			Lower-Bound	Upper-Bound				Lower-Bound	Upper-Bound
EXPERIMENT 4A									
CONCRETE									
1	1.83	1.35	1.60	2.05	1	1.89	1.61	1.58	2.20
2	2.36	1.38	2.12	2.59	2	2.29	1.39	2.02	2.55
3	3.43	1.34	3.21	3.66	3	3.23	1.36	2.97	3.49
4	4.59	1.65	4.32	4.87	4	4.16	1.79	3.82	4.50
5	5.26	1.75	4.97	5.56	5	5.00	1.93	4.63	5.37
ABSTRACT									
1	1.82	1.29	1.59	2.05	1	2.35	1.91	2.02	2.68
2	2.25	1.20	2.03	2.46	2	2.74	1.75	2.43	3.04
3	3.52	1.32	3.29	3.76	3	3.43	1.42	3.18	3.67
4	4.84	1.61	4.56	5.13	4	4.18	1.74	3.87	4.48
5	5.35	1.75	5.04	5.66	5	4.93	2.07	4.57	5.29
EXPERIMENT 4B									
1	2.06	1.57	1.88	2.24	1	2.60	2.02	2.37	2.83
2	2.24	1.43	2.08	2.40	2	3.02	1.86	2.81	3.23
3	3.33	1.52	3.16	3.51	3	3.51	1.62	3.33	3.69
4	4.60	1.66	4.42	4.79	4	4.27	2.02	4.04	4.50
5	5.11	1.81	4.90	5.31	5	4.83	2.13	4.59	5.08
EXPERIMENT 5									
1	1.45	1.12	1.32	1.58	1	1.82	1.50	1.65	1.99
2	1.87	1.30	1.72	2.02	2	2.33	1.38	2.17	2.48
3	3.22	1.36	3.07	3.38	3	3.34	1.23	3.20	3.48
4	4.73	1.75	4.53	4.93	4	4.60	1.90	4.38	4.81
5	5.29	1.99	5.06	5.52	5	5.24	2.03	5.01	5.47

MEAN AND CONFIDENCE INTERVAL GRAPHS

Each graph plots the mean evaluation +/- the 95% confidence interval for the consistent and inconsistent rating polarity conditions. (*BIB = bigger-is-better; SIB = smaller-is-better*)





WEB APPENDIX D: RULING OUT ALTERNATIVE EXPLANATIONS

MEANS AND STDV OF POSSIBLE ALTERNATIVE EXPLANATIONS

Measure	BIB		SIB	
	Mean	SD	Mean	SD
PRETEST				
Mood ¹	63.86	14.62	60.72	17.80
Involvement ¹	66.32	14.06	64.33	17.40
NFC	3.70	0.46	3.64	0.57
EXPERIMENT 1				
Mood ¹	70.40	17.09	70.95	16.51
NFC	3.55	0.54	3.58	0.51
Good Product				
Mood ¹	72.14	15.52	68.70	15.29
NFC	3.58	0.51	3.52	0.44
Bad Product				
Mood ¹	68.57	18.63	73.01	17.46
NFC	3.51	0.57	3.63	0.57
EXPERIMENT 2A				
Mood ²	5.81	1.12	5.43	1.27
Involvement	4.61	0.50	4.55	0.66
EXPERIMENT 2B				
Mood ²	5.74	1.40	5.40	1.38
Involvement	4.60	0.51	4.67	0.48
EXPERIMENT 3				
Mood	3.99	0.98	4.13	0.75
Involvement	4.14	0.74	4.32	0.63
EXPERIMENT 4A				
Mood ²	4.01	0.85	4.06	0.74
Involvement	4.54	0.62	4.53	0.49
Concrete Mindset				
Mood ²	4.01	0.74	4.03	0.72
Involvement	4.45	0.74	4.52	0.51
Abstract Mindset				
Mood ²	4.01	0.98	4.09	0.76
Involvement	4.63	0.44	4.54	0.49
EXPERIMENT 4B				
Mood ²	4.62	0.94	4.75	0.98
Involvement	3.93	0.60	3.94	0.69
EXPERIMENT 5				
Mood ²	4.33	1.68	4.54	1.81
Involvement	4.82	0.38	4.79	0.39

¹100-point scale, ²7-point scale, all others 5-point scale

ANALYSIS DETAILS RULING OUT POSSIBLE ALTERNATIVE EXPLANATIONS

Experiments 2a & 2b

Mood: We averaged mood measures into an index (2a: $\alpha = .97$; 2b: $\alpha = .97$) and conducted an ANOVA with rating polarity as the dependent measure. This analysis for experiment 2a revealed a marginally significant effect of rating polarity where those in the consistent rating polarity reported slightly more positive mood ($M = 5.75$) than those in the inconsistent rating polarity condition ($M = 5.40$, $F(1, 205) = 3.26$, $p = .07$). In order to rule out the possibility that mood could account for the results, we conducted the analysis using mood as an additional independent variable. The effect of mood was not significant ($\beta = .11$, $p = .52$) and the rating polarity x quality level interaction remained significant ($\beta = -.46$, $p = .01$), ruling out the possibility that mood might account for the results. This analysis for experiment 2b revealed a significant effect of rating polarity where those in the consistent rating polarity reported slightly more positive mood ($M = 5.82$) than those in the inconsistent rating polarity condition ($M = 5.43$, $F(1, 216) = 5.59$, $p = .02$). In order to rule out the possibility that mood could account for the results, we conducted the willingness-to-pay analysis using mood as an additional independent variable. The effect of mood was not significant ($\beta = .07$, $p = .92$) and the rating polarity x quality level interaction remained significant ($\beta = -.24$, $p < .01$), ruling out the possibility that mood might account for the results. Involvement: Similarly, to rule out the possibility that participants might have been more involved when attempting to use a rating format with inconsistent rating polarity, we averaged the measures of involvement into an index (2a: $\alpha = .92$; 2b: $\alpha = .95$), and an ANOVA revealed no difference in involvement by condition (2a: $p = .30$; 2b: $p = .44$).

Experiment 3

Mood: We averaged three mood measures into an index ($\alpha = .93$) and a PROC GLM analysis that revealed no difference in mood by condition ($p = .46$). Involvement: Similarly, we averaged the measures of involvement into an index ($\alpha = .90$), and the same PROC GLM analysis revealed no difference in involvement by condition ($p = .22$).

Experiment 4a

Mood: The three mood measures were averaged together ($\alpha = .90$). A regression with rating polarity, mindset, and their interaction as independent measures and the averaged mood measure as the dependent measure confirmed that there was no significant effect of rating polarity, mindset, or their interaction on mood (all $ps > .68$). Involvement: Similarly, the four measures of participant involvement were averaged into an index ($\alpha = .93$), and a regression with the same independent measures and the involvement index as a dependent measure confirmed there was no significant effect of rating polarity, mindset, or their interaction on involvement and attention to the evaluation task (all $ps > .23$).

Experiment 4b

Mood: The mood measures were averaged together ($\alpha = .82$). A regression with rating polarity, mindset, and their interaction as independent measures and the averaged mood measure as the dependent measure confirmed that there was no significant effect of rating polarity, mindset, or their interaction on mood (all $ps > .29$). Involvement: Similarly, the four measures of participant involvement were averaged into an index ($\alpha = .90$), and a regression with the same independent measures and the involvement index as the dependent measure confirmed there was no significant effect of rating polarity, mindset, or their interaction on involvement and attention

to the evaluation task (all $ps > .79$).

Experiment 5

Mood: The three mood measures were averaged together ($\alpha = .98$). A regression with rating polarity, memory type, and their interaction as independent measures and the averaged mood measure as the dependent measure confirmed that there was no significant effect of rating polarity, memory, or their interaction on mood (all $ps > .40$). Involvement: Similarly, the measures of participant involvement were averaged into an index ($\alpha = .78$), and a regression with the same independent measures and the involvement index as a dependent measure confirmed there was no significant effect of rating polarity, memory type, or their interaction on involvement with the evaluation task (all $ps > .35$).

WEB APPENDIX E: ANALYSIS WITH REACTION TIME AS COVARIATE

Analysis for each experiment was conducted as a regression (paradigm pretest, experiments 1) or repeated measures regression (remaining experiments) with the evaluation as the dependent measure and log reaction time as a covariate. Rating polarity was dummy coded (0 = consistent rating polarity, 1 = inconsistent rating polarity).

	<i>Effect</i>	β	<i>SE</i>	<i>DF</i>	<i>t / Z</i>	<i>p</i>
Pretest	Rating Polarity	-1.77	0.77	69	-2.30	0.02
	Reaction Time	0.86	0.47	69	1.84	0.07
Experiment 1	Rating Polarity	4.56	3.42	164	1.33	0.19
	Quality Level	9.70	3.51	164	2.77	0.01
	Rating Polarity x Quality Level	-10.00	4.82	164	-2.07	0.04
	Reaction Time	-1.44	2.79	164	-0.52	0.61
Experiment 2a	Rating Polarity	0.70	0.41	202	1.69	0.09
	Quality Level	1.47	0.08	3095	17.67	<0.01
	Rating Polarity x Quality Level	-0.33	0.12	3095	-2.70	<0.01
	Reaction Time	0.04	0.08	3095	-0.52	0.61
Experiment 2b	Rating Polarity	0.35	0.11	215	3.27	<0.01
	Quality Level	0.54	0.02	3011	34.24	<0.01
	Rating Polarity x Quality Level	-0.10	0.02	3011	-4.53	<0.01
	Reaction Time	0.05	0.03	3011	1.69	0.09
Experiment 3	Rating Polarity	0.87	0.27	87	3.22	<0.01
	Quality Level	0.59	0.04	1243	14.78	<0.01
	Rating Polarity x Quality Level	-0.28	0.06	1243	-4.95	<0.01
	Reaction Time	0.19	0.08	1243	2.53	0.01
Experiment 4a	Rating Polarity	0.14	0.25	162	0.55	0.58
	Quality Level	0.91	0.04	2319	23.06	<.01
	Rating Polarity x Quality Level	-0.11	0.06	2319	-1.76	0.08
	Construal Mindset	-0.11	0.24	162	-0.46	0.64
	Rating Polarity x Construal Mindset	0.75	0.35	162	2.17	0.03
	Quality Level x Construal Mindset	0.06	0.06	2319	0.97	0.33
	3-way interaction	-0.20	0.08	2319	-2.40	0.02
	Reaction Time	-0.04	0.05	2319	-0.85	0.40
Experiment 4b	Rating Polarity	1.02	0.17	197	5.97	<.01
	Quality Level	0.84	0.03	2800	27.88	<.01
	Rating Polarity x Quality Level	-0.28	0.04	2800	-6.49	<.01
	Construal Mindset	-0.06	0.03	197	-1.99	0.05
	Rating Polarity x Construal Mindset	0.14	0.04	197	3.71	0.00
	Quality Level x Construal Mindset	0.01	0.01	2800	1.53	0.13
	3-way interaction	-0.03	0.01	2800	-2.66	0.01
	Reaction Time	-0.05	0.04	2800	-1.18	0.24
Experiment 5	Rating Polarity	0.61	0.16	192	3.76	0.00
	Quality Level	1.06	0.03	2719	40.24	<.01
	Rating Polarity x Quality Level	-0.15	0.04	2719	-3.94	<.01
	Implicit Memory	-0.03	0.02	192	-1.90	0.06
	Rating Polarity x Implicit Memory	0.01	0.03	192	0.59	0.55
	Quality Level x Implicit Memory	0.01	0.00	2719	1.53	0.13
	3-way interaction	-0.01	0.01	2719	-1.59	0.11
	Reaction Time	0.02	0.04	2719	0.63	0.53

WEB APPENDIX F: ORDER EFFECTS ANALYSES

Rating Polarity x Quality Level x Order Effects are highlighted in gray.
Order was coded as 1 to 15 for each of the 15 products in experiments 2a, 2b, 3, 4a, 4b, and 5 and mean centered for the analysis.

		β	<i>SE</i>	<i>DF</i>	<i>t</i>	<i>p</i>
Exp 2a	Rating Polarity	0.74	0.43	204	1.73	0.08
	Quality Level	1.49	0.09	3095	16.82	<0.01
	Rating Polarity x Quality Level	-0.34	0.13	3095	-2.70	<0.01
	Order	-0.18	0.06	3095	-2.95	<0.01
	Rating Polarity x Order	-0.01	0.09	3095	-0.17	0.87
	Quality Level x Order	0.06	0.02	3095	3.29	<0.01
	3-way Interaction	-0.01	0.03	3095	-0.31	0.76
Exp 2b	Rating Polarity	0.36	0.11	217	3.40	<0.01
	Quality Level	0.54	0.02	3060	34.28	<0.01
	Rating Polarity x Quality Level	-0.10	0.02	3060	-4.48	<0.01
	Order	0.0004	0.01	3060	0.03	0.97
	Rating Polarity x Order	-0.02	0.02	3060	-0.94	0.35
	Quality Level x Order	0.001	0.00	3060	0.15	0.88
	3-way Interaction	0.003	0.01	3060	0.66	0.51
Exp 3	Rating Polarity	0.84	0.27	87	3.06	<0.01
	Quality Level	0.59	0.04	1240	14.59	<0.01
	Rating Polarity x Quality Level	-0.28	0.06	1240	-4.88	<0.01
	Order	-0.05	0.03	1240	-1.54	0.12
	Rating Polarity x Order	0.02	0.05	1240	0.47	0.64
	Quality Level x Order	0.02	0.01	1240	1.70	0.09
	3-way Interaction	-0.01	0.01	1240	-0.76	0.45
Exp 4a	Rating Polarity	0.88	0.24	163	3.64	<0.01
	Quality Level	0.96	0.04	2326	23.07	<0.01
	Rating Polarity x Quality Level	-0.30	0.06	2326	-5.14	<0.01
	Mindset	0.08	0.24	163	0.35	0.72
	Rating Polarity x Mindset	-0.68	0.35	163	-1.96	0.05
	Quality Level x Mindset	-0.05	0.06	2326	-0.87	0.39
	Rating Polarity x Quality Level x Mindset	0.17	0.08	2326	2.06	0.04
	Order	-0.03	0.03	2326	-0.86	0.39
	Rating Polarity x Order	0.01	0.04	2326	0.25	0.80
	Quality Level x Order	0.01	0.01	2326	1.39	0.16
	Rating Polarity x Quality Level x Order	-0.01	0.01	2326	-0.77	0.44
	Mindset x Order	0.06	0.04	2326	1.38	0.17
	Rating Polarity x Mindset x Order	-0.16	0.07	2326	-2.52	0.01
	Quality Level x Mindset x Order	-0.02	0.01	2326	-1.28	0.20
Rating Polarity x Quality Level x Mindset x Order	0.06	0.02	2326	2.88	<0.01	

WEB APPENDIX F: ORDER EFFECTS ANALYSES

		β	<i>SE</i>	<i>DF</i>	<i>t</i>	<i>p</i>
Exp 4b	Rating Polarity	1.00	0.17	197	5.89	<0.01
	Quality Level	0.83	0.03	2802	27.42	<0.01
	Rating Polarity x Quality Level	-0.27	0.04	2802	-6.37	<0.01
	Mindset	-0.05	0.03	197	-1.92	0.06
	Rating Polarity x Mindset	0.14	0.04	197	3.70	<0.01
	Quality Level x Mindset	0.01	0.01	2802	1.37	0.17
	Rating Polarity x Quality Level x Mindset	-0.03	0.01	2802	-2.65	<0.01
	Order	-0.04	0.02	2802	-1.78	0.08
	Rating Polarity x Order	0.03	0.03	2802	0.92	0.36
	Quality Level x Order	0.03	0.01	2802	3.97	<0.01
	Rating Polarity x Quality Level x Order	-0.01	0.01	2802	-1.29	0.20
	Mindset x Order	0.002	0.01	2802	0.38	0.71
	Rating Polarity x Mindset x Order	-0.01	0.01	2802	-0.69	0.49
	Quality Level x Mindset x Order	-0.002	0.002	2802	-1.12	0.26
	Rating Polarity x Quality Level x Mindset x Order	0.004	0.002	2802	1.62	0.11
Exp 5	Rating Polarity	0.53	0.17	193	3.14	<0.01
	Quality Level	1.04	0.03	2549	36.48	<0.01
	Rating Polarity x Quality Level	-0.12	0.04	2549	-2.95	<0.01
	Implicit Memory	-0.05	0.02	193	-2.78	<0.01
	Rating Polarity x Implicit Memory	0.02	0.02	193	0.95	0.35
	Quality Level x Implicit Memory	0.01	0.00	2549	2.40	0.02
	Rating Polarity x Quality Level x Implicit Memory	-0.01	0.01	2549	-1.71	0.09
	Order	-0.01	0.02	2549	-0.27	0.78
	Rating Polarity x Order	-0.03	0.03	2549	-1.01	0.31
	Quality Level x Order	0.01	0.01	2549	1.75	0.08
	Rating Polarity x Quality Level x Order	-0.001	0.01	2549	-0.14	0.89
	Implicit Memory x Order	0.004	0.003	2549	1.19	0.23
	Rating Polarity x Implicit Memory x Order	0.001	0.005	2549	0.24	0.81
	Quality Level x Implicit Memory x Order	-0.001	0.001	2549	-0.80	0.42
	Rating Polarity x Quality Level x Judge x Order	-0.0005	0.001	2549	-0.32	0.75
CONTRASTS						
	High Reliance on Implicit Memory: Rating Polarity x Quality Level	-0.19	0.06	2549	-3.29	<0.01
	Low Reliance on Implicit Memory: Rating Polarity x Quality Level	-0.05	0.06	2549	-0.87	0.39

WEB APPENDIX G: EXPERIMENT 4B FLOODLIGHT ANALYSIS

DIFFERENCE IN THE RATING POLARITY x QUALITY LEVEL INTERACTION ACROSS
BIF VALUES

BIF Value (0=most concrete, 24=most abstract)	Interference from Rating Polarity Effect (Rating Polarity x Quality Level Interaction Coefficient)	Lower 95% Confidence Interval	Upper 95% Confidence Interval	<i>t</i> (2810)	<i>p</i>
0	0.02	-0.22	0.26	0.13	0.89
2	-0.03	-0.24	0.17	-0.31	0.76
4	-0.08	-0.25	0.09	-0.91	0.36
6	-0.13	-0.27	0.01	-1.79	0.07
8	-0.18	-0.29	-0.06	-3.05	0.002
10	-0.23	-0.32	-0.13	-4.76	<.0001
12	-0.27	-0.36	-0.19	-6.41	<.0001
14	-0.32	-0.41	-0.23	-7.00	<.0001
16	-0.37	-0.48	-0.26	-6.62	<.0001
18	-0.42	-0.56	-0.28	-6.01	<.0001
20	-0.47	-0.64	-0.30	-5.47	<.0001
22	-0.52	-0.72	-0.32	-5.05	<.0001
24	-0.56	-0.80	-0.33	-4.72	<.0001

Floodlight analysis follows procedures outlined in Spiller et al. (2013).