

UCLA

UCLA Electronic Theses and Dissertations

Title

Known Unknowns in Judgment and Choice

Permalink

<https://escholarship.org/uc/item/0hj5c9s5>

Author

Walters, Daniel

Publication Date

2017

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Known Unknowns in Judgment and Choice

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy
in Management

by

Daniel James Walters

2017

© Copyright by

Daniel James Walters

2017

ABSTRACT OF THE DISSERTATION

Known Unknowns in Judgment and Choice

by

Daniel James Walters

Doctor of Philosophy in Management

University of California, Los Angeles, 2017

Professor Craig R. Fox, Chair

This dissertation investigates how people make inferences about missing information. Whereas most prior literature focuses on how people process known information, I show that the extent to which people make inferences about missing information impacts judgments and choices. Specifically, I investigate how (1) awareness of *known unknowns* affects overconfidence in judgment in Chapter 1, (2) beliefs about the *knowability of unknowns* impacts investment strategies in Chapter 2, and (3) inferences about *forgotten unknowns* influence choices from memory in Chapter 3.

Chapter 1 investigates how overconfidence can stem from neglecting to consider missing information. Most prior research has attributed overconfidence to people focusing disproportionately on evidence favoring the chosen hypothesis relative to its alternatives. In this

chapter, I find that neglecting *unknown* evidence independently contributes to overconfidence. In a first study, respondents answered questions such as, “Which of these fast food items has more calories, a Subway sandwich, or a McDonald's cheeseburger? / How confident are you?” Using a process tracing technique, I found that participants who considered more missing evidence were less overconfident than those that thought about more known evidence. Meanwhile, participants who considered more unknown information answered the same number of questions correctly, resulting in better calibration. In two additional studies, I prompted participants to list unknowns before assessing confidence in their judgments. This “consider the unknowns” technique reduced overconfidence substantially, and was more effective than the de-biasing technique most often prescribed in the research literature (“consider the alternative”). Importantly, considering the unknowns was selective in its impact: it reduced confidence only in domains where participants were overconfident, but did not affect confidence in domains where participants were well-calibrated or under-confident.

Chapter 2 investigates how inferences about the knowability of missing information impacts investment choices. Recent research has found that people intuitively distinguish *aleatory* uncertainty that is inherently random or stochastic (e.g. What is the probability that a fair coin will land heads?) from *epistemic* uncertainty that is attributed to missing knowledge or information (e.g., Which company had a larger market capitalization at the end of 2015, Google or Apple?). In a series of surveys and experiments involving laypeople, experienced investors, and financial advisors, I found that investors who viewed stock market uncertainty as more epistemic/knowable searched for more stock information and were willing to pay more for financial advice, whereas investors who viewed stock market uncertainty as more aleatory/random diversify more. Similarly, when investors were primed to think about epistemic

uncertainty they are more willing to pay for stock information whereas when they were primed to think about aleatory uncertainty they diversified more when completing an incentive-compatible investment task. Taken together, these studies point to the critical role that the perception of the nature of uncertainty can have on an investor's judgments and choices.

Chapter 3 investigates how inferences about forgotten product features impact consumer choices from memory. Consumers must often make product judgments and choices based on information contained in memory. For instance, a consumer may learn information about the Apple Watch, then choose among smart watches later. The consumer may recall some features clearly, such as the Apple Watch's ability to respond to text messages, but also realize that they have forgotten other features entirely. Whereas most prior research has focused on how consumers evaluate remembered information, this research examines how product judgments and choices are also guided by inferences consumers make about forgotten information. First, I found that consumers overestimated how closely the quality of forgotten features resembled remembered features. As a result, consumers tended to choose a product from memory over an equivalent product that is fully described when the remembered features are more positive and choose a product that is fully described over an equivalent product from memory when the remembered features are more negative.

Taken together, this work expands our understanding of how people make inferences about missing evidence, the nature of uncertainty, and forgotten information. Importantly, this work shows how these inferences impact critical outcomes in the field, such as overconfidence, investment strategy, and product choice.

The dissertation of Daniel James Walters is approved.

Hal Ersner Hershfield

Philip Fernbach

Suzanne Bliven Shu

Margaret Joan Shih

Craig R. Fox, Committee Chair

University of California, Los Angeles

2017

TABLE OF CONTENTS

TABLE OF CONTENTS.....	vi
LIST OF TABLES AND FIGURES.....	viii
ACKNOWLEDGEMENTS	ix
CURRICULUM VITA	xi
Chapter 1: Known Unknowns: A Critical Determinant of Confidence and Calibration	1
ABSTRACT.....	2
STUDY 1	7
STUDY 2	13
STUDY 3	17
GENERAL DISCUSSION	22
Chapter 2: Two Dimensions of Uncertainty Predict Investment Strategy.....	26
ABSTRACT.....	27
STUDY 1	31
STUDY 2	36
STUDY 3	40
STUDY 4	45
STUDY 5	49
GENERAL DISCUSSION	51
Chapter 3: The Influence of Forgotten Features on Product Choice	54
ABSTRACT.....	55
STUDY 1	63
STUDY 2	67
STUDY 3	72
STUDY 4	77
STUDY 5	83
GENERAL DISCUSSION	86
Conclusion	91
Tables.....	93
Figures.....	94
Appendix A: Materials for Studies in Chapter 1	100
Study 1 Two Alternative Forced Choice Questions	100

Study 1 Known unknown rating instructions.....	102
Study 1, 2, and 3 Typical Text Responses	104
Study 2 Performance Incentive	106
Study 2 Questions	107
Study 3 question stems and number of questions.	108
Study 3 Procedure for identifying overconfident domains and calibrated/underconfident domains.	109
Appendix B: Materials for Studies in Chapter 2.....	110
Study 1-5, 6-item EARS Instructions:	110
Study 1 and 2 Exclusion Criteria	111
Study 1 DOSPERT questions	114
Study 3 Primes:	115
Study 3 Stock Information	117
Study 4 Research Report.....	118
Appendix C: Materials for Studies in Chapter 3.....	119
Study 1 Retailers	119
Study 2 Retailers	120
Study 3 Movies	121
Study 4 TV Features	122
Study 5 Negative Reviews	123
References.....	125

LIST OF TABLES AND FIGURES

Table 1. Correlations between Chapter 3, Study 4 measures	93
Figure 1: Overconfidence, Chapter 1, Study 2	94
Figure 2: Effect of Consider the Unknowns, Chapter 1, Study 3	95
Figure 3: Effect of Consider the Unknowns Across Domains, Chapter 1, Study 3.....	96
Figure 4: Effect of prime X financial Literacy Chapter 2, Study 2	97
Figure 5: Mediation of Gift Card Choice, Chapter 3, Study 2.....	98
Figure 6: Mediation of TV Choice, Chapter 3, Study 4.....	99

ACKNOWLEDGEMENTS

This dissertation is dedicated to the family that supported me throughout my PhD and the mentors that fostered my growth as an academic. First, thank you to my incredible wife Kedist for providing me with strength throughout the past years. It is only with you at my side that I have been able to complete this journey and be truly proud of my life as we make the leap into our next chapter. Thank you also to my mother, Ann Walters, who has helped me in so many more ways than she can ever know. She has taught me by example how to be a good parent and how to provide generosity, patience, and love to foster growth. I only hope that I can do the same for my daughter Evan. I also dedicate this dissertation to my late father, Steve Walters, who taught me that success flows from good judgment and the proper application of hard work. I miss you and wish you were here to share this with me, but feel your presence as I become more like you every year.

Thank you also to my colleagues and friends that supported my academic growth. First, thank you to Craig Fox, my incredible advisor for always stoking that fire in my belly. I always came out of our meetings inspired to do research, feeling the weight and importance of our work. Thank you to Phil Fernbach who taught me that an important ingredient to writing a strong paper is breaking down ideas into the simplest, most easily understood constituents. Thank you also to Hal for great advice and friendship. You have a great ability to understand people and always provided me with sound advice for any situation on the job market. Finally, thank you David Tannenbaum for being a great mentor and teaching me most of what I know about analyzing data. Thank you also to Margaret Shih for giving me a shot when I was just an MBA and to

Geoff Ho and Carsten Erner who were generous with their time and helped to mentor me early on during my PhD.

The final step of my journey could not have been made without my wonderful daughter. Evan, thank you for focusing me in a critical time and giving me the motivation to make a great life for our family. You are a ball of joy and a constant source of inspiration.

CURRICULUM VITA

DANIEL WALTERS

EDUCATION

UCLA ANDERSON SCHOOL OF MANAGEMENT <i>PhD Candidate</i> <i>Advisor: Craig R. Fox</i>	Los Angeles, CA Expected June 2017
UCLA ANDERSON SCHOOL OF MANAGEMENT <i>MBA</i> June 2011	Los Angeles, CA
HEC SCHOOL OF MANAGEMENT <i>MBA Study Abroad Program</i> Fall 2010	Paris, France
NEW YORK UNIVERSITY, STERN SCHOOL OF BUSINESS BS Economics & Mathematics	New York, NY June 2006

HONORS & AWARDS

UCLA Dissertation Year Fellowship	June 2016
Dare to Care Fellowship (\$3,000)	June 2016
Distinguishing Two Dimensions of Subjective Uncertainty.	August 2014
National Science Foundation Grant with Craig R. Fox and Gülden Ülkümen Total budget: \$558,000.	
UCLA Anderson Graduate Fellowship (\$200,000)	2011 - 2016
UCLA Anderson MBA Merit Based Fellowship (\$8,000)	October 2010
Charles E. Davidson Fellowship (\$20,000)	May 2009
New York University Stern Academic Scholarship (\$5,000)	August 2005
Occidental College Academic Merit Scholarship (\$40,000)	August 2004

PUBLICATIONS

- Walters, Daniel J., Philip M. Fernbach, Craig R. Fox and Steven A. Sloman (2016), "Known unknowns: A critical determinant of confidence and calibration," *Management Science*, in press.
- Fox, Craig R., Carsten Erner and Daniel J. Walters (2015). "Decision under risk: From the field to the laboratory and back," Chapter 2 of G. Keren and G. Wu (Eds.) *Blackwell Handbook of Judgment and Decision Making: An Interdisciplinary Perspective*, New York: Wiley.
- Ho, Geoff C., Margaret Shih and Daniel J. Walters (2012). "Labels & leaders: The influence of task framing on leadership emergence," *The Leadership Quarterly*, 23 (5), 943-952.

WORKING PAPERS

Walters, Daniel J. and Hal E. Herschfield, “How do we judge what we forget?: Overconfidence when making product choices from memory,” revision invited by *Journal of Marketing Research*.

Walters, Daniel J., Gülden Ülkümen, Carsten Erner, Craig R. Fox and David Tannenbaum, “Two dimensions of unknowns predict investment decisions,” *Target: Journal of Marketing Research*.

Walters, Daniel J., Carsten Erner, Craig R. Fox, Marc Scholten and Daniel Read, “Debt Aversion: Anomalous in the lab, advantageous in practice,” *Target: Journal of Consumer Research*.

Scholten, Marc, Daniel J. Walters, Carsten Erner, Craig R. Fox and Daniel Read, “The Debt Battery.”

INVITED TALKS AND CONFERENCE PRESENTATIONS

Walters, Daniel J., Carsten Erner, Craig R. Fox, Marc Scholten and Daniel Read, “Debt Aversion: Anomalous in the lab, advantageous in practice,” *Center for Research on Consumer Financial Decision Making*, University of Colorado, Boulder, CO, 2016

Walters, Daniel J., Craig R. Fox, Daniel Read and Marc Scholten, “Loss aversion for time and money: Reference dependent intertemporal choices” *Society for Consumer Psychology*, Phoenix, AZ, 2015

Walters, Daniel J., Philip M. Fernbach, Craig R. Fox and Steven A. Sloman, “Known unknowns determine confidence and calibration,” *Society for Judgment and Decision Making*, Long Beach, CA, 2014

Walters, Daniel J., Philip M. Fernbach, Craig R. Fox and Steven A. Sloman, “Known unknowns determine confidence and calibration,” *Behavioral Decision Research in Management*, London, UK, 2014

Walters, Daniel J., Philip M. Fernbach, Craig R. Fox and Steven A. Sloman, “Emerging evidence from the science of intellectual humility across the lifespan,” *Association of Psychology Science*, San Francisco, CA, 2014

Walters, Daniel J., Philip M. Fernbach, Craig R. Fox and Steven A. Sloman, “Known unknowns determine confidence and calibration,” *Society for Consumer Psychology*, Miami, FL, 2014

Walters, Daniel J., Philip M. Fernbach, Craig R. Fox and Steven A. Sloman, “Known unknowns determine confidence and calibration,” *Subjective Probability, Utility, and Decision Making Conference*, Barcelona, Spain, 2013

Chapter 1: Known Unknowns: A Critical Determinant of Confidence and Calibration

ABSTRACT

We propose that an important determinant of judged confidence is the evaluation of evidence that is unknown or missing, and overconfidence is often driven by the neglect of unknowns. We contrast this account with prior research suggesting that overconfidence is due to biased processing of known evidence in favor of a focal hypothesis. In Study 1, we asked participants to list their thoughts as they answered two-alternative-forced choice trivia questions and judge the probability that their answers were correct. Participants who thought more about unknowns were less overconfident. In Studies 2 and 3 we asked participants to list unknowns before assessing their confidence. “Considering the unknowns” reduced overconfidence substantially, and was more effective than the classic “consider the alternative” debiasing technique. Moreover, considering the unknowns selectively reduced confidence in domains where participants were overconfident, but did not affect confidence in domains where participants were well-calibrated or underconfident.

In the run-up to the Iraq war of 2003 many leaders in the United States expressed great confidence that the Iraqi President, Saddam Hussein was developing weapons of mass destruction (WMDs). In a letter sent to President George W. Bush in 2001, ten of the most influential congressman, both Democrats and Republicans, wrote "There is no doubt that ... Saddam Hussein has reinvigorated his weapons program. Reports indicate that biological, chemical and nuclear programs continue apace and may be back to pre-Gulf War status."¹ Senator Jay Rockefeller expressed the same sentiment in a 2002 speech: "There is unmistakable evidence that Saddam Hussein is working aggressively to develop nuclear weapons and will likely have nuclear weapons within the next five years."² As we now know, there were no WMDs, so these statements expressing "no doubt" and "unmistakable evidence" apparently reflected overconfidence that had major geopolitical consequences. While this example may be extreme it is not unusual. Overconfidence has been implicated in a wide range of decision errors, from going to war (Johnson 2004) to treatment of medical conditions (Baumann, Deber, and Thompson 1991; Oskamp 1965) to corporate investments (Malmendier and Tate 2005) to market entry (Camerer and Lovallo 1999; Mahajan 1992).

A great deal of research has attempted to understand the sources of error in judging confidence with an eye to developing debiasing techniques. Much of this research has attributed overconfidence to a systematic tendency to seek or overweight known evidence for a favored hypothesis over its alternatives. In the case of the Iraq war, overconfidence may have been driven in part by the Bush administration promoting the hypothesis that Iraq was developing

¹ Letter to President Bush, Signed by Senator Bob Graham and others, December 5, 2001

² Senator Jay Rockefeller, October 10, 2002

WMDs and the bias among observers to seek and overweight evidence confirming this hypothesis. An abundance of research has found that people tend to focus disproportionately on evidence for a focal hypothesis relative to alternatives (Hoch 1985; Klayman 1995; Koriat, Lichtenstein, and Fischhoff 1980), and they tend to seek evidence consistent with the focal hypothesis as part of a positive test strategy (Klayman and Ha 1987; Mynatt, Doherty, and Tweney 1977; Nickerson 1998), wishful thinking (Babad 1987), motivated reasoning (Kunda 1990), or to protect their self-image from failure and regret (Larrick 1993). One reason this approach to understanding overconfidence has been so influential is because it has led to successful debiasing techniques that tend to improve judgment calibration. Overconfidence can be reduced by prompting people to “consider the alternative” (Koriat, Lichtenstein, and Fischhoff 1980) or by designating a member of a decision-making team to advocate for the alternative (“devil’s advocate technique”; Schwenk and Cosier 1980).

A second class of theories of confidence represents the mapping between balance of known evidence and judged probabilities. Griffin and Tversky (1992) distinguish strength of evidence (i.e., balance) from weight of evidence (i.e., reliability or diagnosticity). They argue that when judging probabilities people tend focus on strength of evidence and give insufficient regard to weight of evidence. This can contribute to both overconfidence (when strength of evidence is high and weight of evidence is low) and underconfidence (when strength of evidence is low and weight of evidence is high). People focus on strength of evidence while neglecting weight of evidence because they overestimate the predictive validity of evidence that is representative (Kahneman and Tversky 1979), internally consistent (Kahneman and Tversky 1973), and based on small samples (Tversky and Kahneman 1971). Similarly, in Support Theory (Rottenstreich and Tversky 1997; Tversky and Koehler 1994), probability is determined by the

perceived balance of evidence for a hypothesis relative to its alternative. Overconfidence can occur due to scaling the perceived balance of evidence to overly extreme judged probabilities (see Fox 1999), for instance when perceived evidence is seen as especially predictive of outcomes (Tannenbaum, Fox, and Ülkümen 2016), or when the environment does not provide particularly diagnostic cues (Brenner, Griffin, and Koehler 2005). In evidence accumulation models confidence is determined by weighting evidence based on feeling (Ferrell and McGoey 1980) or self-consistency (Koriat 2012), and overconfidence can occur when these cues are overestimated.

We propose that when assessing confidence people may also look directly to specific pieces of *unknown evidence* to determine how to weight or scale the balance of known evidence. By unknown evidence, we mean a variable whose value is unknown, but if it were known should change one's level of confidence. For instance, prior to the invasion of Iraq, Saddam Hussein's motivation for not cooperating with weapons inspectors was unknown to most American observers. Mr. Hussein may have wanted the world to believe that he *did possess WMDs* (to increase the perceived strength of the Iraqi military) or that he *did not possess WMDs* (to reduce the likelihood of a US-led invasion). Becoming aware of this important unknown factor would not change the information available to a judge. However, awareness of the unknown is likely to decrease confidence by making the judge aware that he or she is missing critical information. Unknown evidence can potentially support the focal or an alternative hypothesis once the unknown is resolved. So being aware of more unknown evidence should generally lead to less extreme confidence in both outcomes.

Biased evaluation of known evidence clearly plays a role in overconfidence, but failure to adequately consider unknowns may be equally important. A growing body of literature shows

that people tend to think the world is simpler and more predictable than it is because they focus on what they know and tend to neglect what they do not know. For instance, people tend to think they understand various types of causal systems, from machines to public policies, in much greater detail than they actually do (Alter, Oppenheimer, and Zemla 2010; Fernbach et al. 2013; Rozenblit and Keil 2002). People also tend to neglect unknown causes of system failure when diagnosing problems such as why a car won't start (Fischhoff, Slovic, and Lichtenstein 1978), and they underestimate the possibility of unknown or unexpected delays in the planning fallacy (Buehler, Griffin, and Ross 1994). People also exhibit a 'censorship bias' in which they fail to account for missing sample information when forming beliefs about an underlying population (Feiler, Tong, and Larrick 2013). Similarly, consumers tend to neglect unknown or unmentioned attributes when evaluating products (Sanbonmatsu, Kardes, and Herr 1992; Sanbonmatsu, Kardes, and Sansone 1991). More generally, Kahneman (2011) uses the focus on known relative to unknown information as an organizing principle for many phenomena in judgment and decision-making, that he refers to as the 'What You See is All There Is' (WYSIATI) principle.

We have proposed that judged confidence depends in part on the judge's assessment of how much evidence is missing or unknown. We predict that greater appreciation of unknowns will be associated with judged probabilities that tend more towards the "ignorance prior" probability of $1/n$ in an n -alternative forced choice paradigm (e.g., $1/2$ when there are two alternatives) whereas less appreciation of unknowns will be associated with more extreme confidence judgments that depart more from the ignorance prior. Consistent with this hypothesis, previous studies suggest that when people are less knowledgeable they provide less extreme probability judgments. Fox and Clemen (2005) report that judged probabilities of n exclusive and exhaustive events—for example, the branches from a chance node in a decision tree—were

biased more strongly toward probabilities of $1/n$ for events about which participants had less knowledge or expertise. Likewise, See, Fox, and Rottenstreich (2006) found that judged probabilities were biased more strongly toward $1/n$ when participants had less opportunity to learn the frequencies of observed events or when they reported feeling less confident in what they had learned.

In Study 1 we use a correlational, thought listing paradigm to test whether differences in consideration of unknowns predict differences in confidence and overconfidence, controlling for the balance of known evidence. We also examine whether under-appreciation of unknowns is associated with overconfidence. We predict that prompting people to consider unknowns will reduce overconfidence. In Studies 2 and 3 we introduce a novel debiasing technique, “consider the unknowns,” in which participants are asked to reflect on what they do not know before reporting their confidence, and we compare the efficacy of this technique to the classic “consider the alternative” intervention (Koriat, Lichtenstein, and Fischhoff 1980).

STUDY 1

We asked participants to judge the probability of making a correct choice in a two-alternative forced choice (2AFC) task involving general knowledge questions. The 2AFC paradigm is a well-studied context in which people often exhibit overconfidence (for reviews see (Griffin and Brenner 2004; Koehler, Brenner, and Griffin 2002; McClelland and Bolger 1994). As participants completed the task we also asked them to provide reasons for their judgments using a thought listing procedure (Johnson, Häubl, and Keinan 2007). We then asked participants to self-code each of their reason on the extent to which it referred to known versus unknown evidence. In addition, we asked two hypothesis-blind judges to code the extent to which each

reason supported the chosen or alternative option. We predicted that respondents would exhibit lower confidence to the extent that they thought about more unknown evidence and that this relationship would hold after controlling for the balance of known evidence.

Methods

We recruited 134 students at the University of Colorado Boulder to participate in a laboratory experiment in exchange for a \$3 payment (49% female; mean age = 20.0). We first asked them to answer ten 2AFC questions, each with two possible answers adapted from Klayman et al. (1999); a complete set of questions is provided in Appendix A. After answering each question, we asked participants to report their confidence by estimating the probability that they correctly answered the question, on a scale from 50% to 100%.

For the first 3 of 10 questions (questions 1-3 in Appendix A) we asked participants to list the reasons for their confidence:

As you answer the question, please think of all the reasons that make you {more/less} confident you know the answer and all the reasons that make you {less/more} confident. We will ask you to enter your reasons one at a time. Type your first complete reason in the box below and, as soon as you are done, hit the “enter” key to submit it. You may enter your reasons in any order.”

The order of the words ‘more’ and ‘less’ was randomly determined for each participant and had no effect on confidence or answer choice. Participants could list as many or as few reasons as came to mind. The entered reasons then appeared and participants had an opportunity to enter more reasons. Participants listed reasons while viewing the 2AFC question and they could change both their answer and confidence while listing reasons.

After completing all ten questions, we reminded participants of each of the reasons they provided for the first three questions. We then asked them to rate each reason as being about known or unknown evidence on a 1-7 scale (1=completely known; 7=completely unknown). We explicitly asked participants to rate how known versus unknown the reason was rather than how much each reason improved the participant's estimate in order to make sure we were measuring the content of the reason, rather than the effect of the reason on confidence. A sample of the rating instructions can be found in Appendix A. Finally, we collected demographic data and debriefed participants.

Results

Unknown Rating and Reasons Generated. For the three question for which participants provided and rated reasons for their confidence estimates, participants provided an average of 2.36 reasons per question with an interquartile range of (2.35, 2.56). We calculated participants' average rating of reasons for how much they involved unknown evidence (1 = completely known; 7 = completely unknown). The mean rating was 3.45 with an interquartile range of (2.56, 5.33), and 63% of participants had an average rating below the scale midpoint, suggesting that most participants reported more known than unknown evidence. Reasons rated as known tended to be statements of facts whereas reasons rated as unknown tended to be statements about missing information or lack of relevant knowledge. Appendix A provides examples of representative known and unknown reasons generated by participants.

Confidence, Percent Correct, and Overconfidence. Across the three questions where reasons were provided mean confidence ratings were 67.4% while on average participants answered 62.2% of questions correctly. For each participant, we calculated overconfidence

following conventional methods (see Griffin and Brenner 2004; Koehler, Brenner, and Griffin 2002; McClelland and Bolger 1994) by subtracting the percentage of all items answered correctly from average confidence, resulting in mean overconfidence of 5.2%, significantly more than 0%, $t(133) = 2.36, p < .05$, replicating previous work (e.g., Koehler, Brenner, and Griffin 2002). Confidence, percent correct and overconfidence did not vary significantly for the seven questions where no reasons were provided compared to the three where reasons were provided.

We next examined the relationship between unknown ratings, confidence, percent correct and overconfidence across the three questions for which participants provided reasons. We first calculated the average confidence and percent correct on these question. We regressed the average confidence judgment on the average unknown rating. As we predicted, participants who provided reasons that they rated as more unknown were less confident, $b = -3.11, 95\% \text{ CI } [-4.63, -1.59], p < .001$. We also regressed percent correct on the known vs. unknown rating and found no significant relationship, $b = 0.66, 95\% \text{ CI } [-2.79, 4.11], p > .5$. We then regressed overconfidence on unknown ratings. Participants who generated reasons that they rated as more unknown exhibited less overconfidence $b = -3.77, 95\% \text{ CI } [-7.38, -0.16], p < .05$. To assess the level of unknown rating at which overconfidence becomes significant we conducted a floodlight analysis (Spiller et al. 2013). The Johnson Neyman point occurred at an unknown rating of 3.1, meaning that at this level of average unknown rating and above it, overconfidence did not significantly differ from 0. Below this average unknown rating, participants were significantly overconfident. At no level of average unknown rating were participants underconfident.

Balance of Known Evidence. We asked two hypothesis-blind coders to code participants' reasons according to the extent to which they appear to support the chosen vs. alternative option, using a 1 to 7 scale (1 = strong support of alternative option; 7 = strong support of the chosen

option). Coders were not provided with the unknown rating or any other data besides the study questions and participant reasons. Nine participants did not provide reasons on at least one of the questions and were not scored by coders. Inter-rater reliability of these scores was high (Cronbach's $\alpha = .80$). Not surprisingly, mean balance of known evidence was 5.33 in favor of the chosen option, with an interquartile range of (4.81, 5.58). Appendix A provides examples of representative reasons coded as supporting the chosen and the alternative options. Rated support was not significantly correlated with unknown rating, ($r = -.12, p = .201$). Focusing only on the questions where participants provided and self-coded reasons, we ran three separate regressions with balance of known evidence as the independent variable and either confidence, percent correct or overconfidence as the dependent variable. Participants who provided reasons that were rated as more supportive of the focal compared to alternative hypothesis were marginally more confident in their choices $b = 2.85, 95\% \text{ CI } [-0.53, 6.24], p = .098$. Balance of known evidence did not significantly predict percent correct, $p > .1$, or overconfidence, $p > .5$.

We next conducted hierarchical regressions with average confidence across the three questions for which participants provided reasons as the dependent variable, and known versus unknown rating and balance of known evidence as the predictors. The model R-squared increased from .02 to .15 when adding known vs. unknown rating to balance of known evidence, $F(1,122) = 18.15, p < .0001$. When adding balance of known evidence to known versus unknown rating, the R-squared marginally increased, from .11 to .15, $F(1,122) = 3.68, p = .057$. This is consistent with our hypothesis that known unknowns contribute to confidence in addition to the balance of known evidence for the chosen versus alternative option.

Within-participants analysis. Because each participant rated multiple items, we were also able to perform a within-participant analysis to examine if an individual's confidence, percent

correct and/or overconfidence varied as he or she listed reasons that were more unknown across different questions. For each participant, we examined the relationship between question-level known versus unknown rating and confidence, accuracy, and overconfidence. For each of the three questions we recorded judged confidence and unknown rating. We scored accuracy as a 1 if correct and a 0 if incorrect, and scored overconfidence as confidence minus accuracy. To analyze the data we used a linear regression with unknown rating for a particular question as the independent variable and confidence as the dependent variable while clustering standard errors by participant. Replicating the between-participant analysis, participants were less confident when they provided more unknown reasons, $b = -3.73$, 95% CI [-4.45, -3.00], $p < .001$. Next, we ran the same regression with overconfidence as the dependent variable. Again replicating the between-participant analysis, higher unknown ratings were related to less overconfidence $b = -6.97$, 95% CI [-9.62, -4.32], $p < .001$. Finally, we ran the same regression with percent correct as the dependent variable. Interestingly, higher unknown ratings significantly predicted percent correct, $b = 3.25$, 95% CI [0.59, 5.90], $p < .05$, a result that we did not predict *ex ante*.

Discussion

This study showed that appreciation of unknowns is related to both confidence and overconfidence. Focusing on more known evidence was associated with greater overconfidence whereas generating reasons that were rated as entailing more unknown evidence was associated with less overconfidence. Previous research has attributed confidence primarily to the processing of the balance of known evidence. Unknown ratings significantly predicted confidence after controlling for the balance of known evidence, suggesting that consideration of unknowns also contributes to judged confidence.

While the results of Study 1 support our hypothesis concerning the role of known unknowns, we acknowledge that the evidence is correlational and thus open to alternative interpretations. For instance, it is possible that those who felt less confident were more likely to reference unknowns rather than the other way around. In Studies 2 and 3 we experimentally manipulate consideration of unknowns to provide causal evidence of the determinants of overconfidence.

STUDY 2

In Study 2 we manipulate thinking about unknowns by explicitly asking some participants to “consider the unknowns” (CTU) and we compared the effectiveness of this intervention to the classic “consider the alternative” (CTA) debiasing intervention, in which people are asked to consider known evidence for the alternative hypothesis (Koriat, Lichtenstein, and Fischhoff 1980). Considering the alternative has been shown to reduce overconfidence, in part by increasing the percent correct. For example, Koriat, Lichtenstein, and Fischhoff found that percent correct in the control condition was 62.9% compared to 69.7% when people were asked to consider the alternative in the 2AFC paradigm. We believe that as people consider the alternative they sometimes correctly realize that there is more evidence in favor of the alternative and switch their choice. Thus, considering the alternative can increase percent correct and decrease confidence. In contrast, considering the unknowns should reduce overconfidence only by reducing misplaced confidence, and should not cause people to switch their choice.

Methods

We recruited 254 participants at the University of California, Los Angeles from an online university subject pool to participate in a laboratory experiment in exchange for \$3 dollars plus a

performance incentive (75.7% female; mean age= 21.0). The performance incentive could range up to \$212 (see Appendix A for details).

Participants assessed their confidence that they provided the correct answer to each of eight general knowledge questions in a four-alternative forced choice (4AFC) format. A complete list of questions is displayed in Appendix A. We randomly assigned participants to one of three conditions: no treatment, consider the alternative, and consider the unknowns. In the no treatment condition participants answered the questions and estimated their confidence without providing any additional information. In the consider the alternative (CTA) condition we adapted the procedure from Koriat, Lichtenstein, and Fischhoff (1980) in which participants in a 2AFC paradigm were prompted to list reasons supporting the non-chosen option (the alternative hypothesis) before making a confidence judgment. In our study, we asked participants to generate reasons supporting one of three possible non-chosen options:

“Write down in the spaces provided two reasons that support one of the alternative choices (non-chosen options). Please write the best reasons you can think of that provides evidence for the options you have rejected. For example, in answering the question: "Which of these cars has a larger engine by volume: Mitsubishi Lancer, Nissan Altima, Mazda CX-5, or Subaru Impreza?" If you chose 'Nissan Altima' you would then list reasons that the correct answer might be the Lancer, the CX-5 or the Impreza.”

In the consider the unknowns (CTU) condition we asked participants to:

“Write down in the space provided two pieces of missing information or two unknown factors that would help you determine the correct choice, if known. For example, in answering the question: "Which of these cars has a larger engine by volume: Mitsubishi

Lancer, Nissan Altima, Mazda CX-5, or Subaru Impreza?" An unknown might be: 'I don't know what a CX-5 is,' or 'I don't know if a Lancer is a sedan or an SUV'. What's important is that you write down two factors that are unknown to you.”

Appendix A shows examples of representative reasons generated by participants in the CTU and CTA conditions.

Results

Figure 1 displays the mean level of confidence, percent correct and overconfidence across the three conditions. Confidence was calculated as the average level of confidence across all eight questions for each participant, percent correct was calculated as the percent correct across all eight questions, and overconfidence was calculated as the difference between the two.

We first analyzed the two treatment conditions against the no treatment condition and against each other. Participants in the CTU condition exhibited lower confidence than those in the no treatment condition, 56.8% vs. 61.1%, $t(170) = 2.14$ $p < .05$ and marginally lower confidence than those in the CTA condition, 61.4%, $t(166) = 1.96$, $p = .052$. Confidence in the consider the alternative condition did not differ significantly from the no treatment condition, $t(166) < 1$, *ns*.

Percent correct in the CTU condition was not significantly different than in the no treatment condition, 41.3% vs. 37.6%, $t(170) = 1.46$, $p > .1$ or the CTA condition, 44.2%, $t(166) = 1.18$, $p > .1$. Percent correct in the CTA condition was significantly higher than the no treatment condition, $t(166) = 2.59$, $p = .01$.

Overconfidence in the CTU condition was significantly lower than in the no treatment condition, 15.5% vs. 23.5%, $t(170) = 2.60, p = .01$ and was no different than in the CTA condition, 17.2%, $t(166) < 1, ns$. Overconfidence in the CTA condition was marginally lower than in the no treatment condition, $t(166) = 1.82, p = .070$.

Discussion

Considering the unknowns reduced confidence, resulting in decreased overconfidence relative to the no treatment condition. In contrast, considering the alternative did not reduce confidence but did improve percent correct, resulting in marginally less overconfidence than the no treatment condition. Thus, both debiasing techniques showed some efficacy, but considering the unknowns was more effective at reducing confidence.

One limitation of Studies 1 and 2 is they do not distinguish whether considering the unknowns generally improves calibration (i.e., meta-knowledge concerning one's accuracy) or whether it merely reduces confidence on questions where people are already overconfident. The downside of a general reduction in confidence is that when people are ordinarily well-calibrated it would lead to underconfidence, and where people are ordinarily underconfident it would exacerbate this bias. Study 3 allows us to examine the extent to which improvements in calibration following the consider the unknowns (CTU) intervention reflect a nonspecific reduction in confidence versus selective adjustment when confidence is misplaced.

STUDY 3

We designed Study 3 to replicate and extend the results of Study 2 by enhancing the design in three respects. First, to address the possible concern that the questions in Study 2 may have been especially difficult, which can lead to overconfidence through unbiased judgment error (Erev, Wallsten, and Budescu 1994; Gigerenzer, Hoffrage, and Kleinbölting 1991; Soll 1996), we randomly generated the questions from a database of 778,169 questions across 9 domains provided to us by Jack Soll (personal communication, November, 2013). Second, in

Study 3 we used a within-participant comparison between control and treatment to generalize the results beyond the between-participant design of Study 2. Finally, to establish the generality of the effects, Study 3 relies on a 2AFC paradigm whereas Study 2 used 4AFC.

Random stimulus sampling and the 2AFC format provide an additional benefit. Because we expect baseline overconfidence to vary across domains (see Klayman et al. 1999), Study 3 allows us to examine the extent to which improvements in calibration due to the consideration of unknowns (CTU) prompts are driven by a general reduction in confidence or selective adjustments that depend on the degree of misplaced confidence. If CTU instead has a selective effect, it can provide a more useful and informative method for reducing confidence. To test this we compare changes in overconfidence in domains where participants are normally overconfident versus those where they are normally well-calibrated or underconfident.

Methods

Participants. We recruited 270 participants through a Qualtrics panel in exchange for \$4 (66.3% female; mean age= 49.2). One participant did not finish the study and nineteen participants (7%) requested that their data not be used in an opt-out option in the study debrief, leaving a sample size of 250.

Stimuli and Procedure. Participants answered twenty general knowledge questions in a 2AFC format and assessed their confidence that they provided the correct answer. For each question, we asked participants to pick the correct answer and assess their confidence on a 50% to 100% scale. The twenty questions were grouped into two blocks of ten questions each: the first block was the no treatment block and the second block was the treatment block. Before the first block, participants read a brief set of instructions, completed a practice problem and then

completed the ten questions with each question presented on a separate screen. Next, participants were randomly assigned to either the consider the alternative (CTA) or consider the unknowns (CTU) treatment condition. Depending on condition, participants read instructions similar to CTA or CTU conditions used in Study 2, and completed the second block of questions, this time elaborating on either the alternative or unknowns for each question, following the procedure of Study 2. Appendix A shows examples of representative reasons generated by participants in the CTU and CTA conditions.

Each participant received a randomly selected sample of questions drawn from a population of 778,169 question combinations developed by Jack Soll and colleagues. A complete list of question domains is displayed in Appendix A. Prior to the study, we created all possible question combinations then we randomly selected five questions per domain, for a total of forty-five questions. Each participant received twenty of these questions, sampled at random without replacement, following a method similar to Klayman et al. (1999)

Results

Figure 2 displays mean confidence, percent correct and overconfidence in the CTU and CTA conditions for the first ten questions (where there was no treatment) and the last ten questions, where participants considered the unknowns or the alternative. Replicating Study 2, considering the unknowns reduced confidence and overconfidence, and in this case, clearly had no effect on percent correct. In line with Study 2, considering the unknowns was more effective at reducing confidence than considering the alternative. For participants in the CTU condition, confidence was lower after generating unknowns than when answering the questions with no treatment, 62.8% vs. 67.8%, $t(120) = 6.14$ $p < .001$. For participants in the CTA condition,

confidence was also slightly lower after generating alternatives than when answering the questions with no treatment, 66.0% vs. 68.0%, $t(128) = 2.37, p < .05$. However, the effect of considering the unknowns on confidence was larger than considering the alternative, $t(248) = 2.55, p = .01$. Considering the unknowns also reduced overconfidence relative to no treatment, from 5.5% to 0.8%, $t(120) = 2.70, p < .01$, whereas considering the alternative did not significantly reduce overconfidence, from 4.5% to 3.4%, $t(128) < 1, p > .5$. While the reduction in overconfidence in the CTU condition (4.7%) was greater than in the CTA condition (1.1%), this difference did not reach statistical significance, $t(248) = 1.38, p = .16$. However, overconfidence was not statistically different from 0 after considering unknowns, $t(134) < 1, p > .5$, whereas after considering the alternative overconfidence persisted, $t(128) = 2.30, p < .05$. Unlike in Study 2 neither manipulation significantly affected percent correct (means for CTU versus no treatment = 62.0% vs. 62.3%, $p > .5$; means for CTA versus no treatment = 62.6% vs. 63.5%, $p > .5$).

We next examined whether considering the unknowns had a larger effect on answers where participants are normally more overconfident. We first identified domains for which participants exhibited statistically significant overconfidence, and domains for which they exhibited calibrated or underconfident judgment. We identified domains using a split-sample method similar to Klayman et al. (1999) so that we could rule out regression to the mean as a trivial mechanism driving improvement (see Appendix A for additional details). Participants were overconfident in four domains (president elected first, food calories, beverage calories, and atomic weight) and calibrated or underconfident in five domains (country life expectancy, distance from Kansas City, state populations, movie box office revenue, and car miles per gallon). For each participant we computed four overconfidence scores: (1) overconfident

domains with a treatment, (2) overconfident domains without a treatment, (3) calibrated/underconfident domains with a treatment, and (4) calibrated/underconfident domains without a treatment (see Figure 3).

We analyzed the CTU and CTA conditions separately using within-participant regression models, with overconfidence as the dependent variable. The independent variables were domain type (overconfident vs. calibrated/underconfident), treatment (treatment vs. no treatment), and their interaction. In overconfident domains, overconfidence was lower after considering the unknowns than when answering the questions with no treatment, 6.8% vs. 15.3%, $b = 8.5$, 95% CI [3.2, 13.9], $p < .01$. In contrast, in calibrated/underconfident domains, considering the unknowns had no significant effect, -3.9% vs. -2.6% $b = 1.3$, 95% CI [-4.0, 6.2], $p > .5$. The interaction between domain type and treatment was marginally significant, indicating that the effect of considering the unknowns was larger in overconfident domains, with a 8.5% reduction in confidence after considering unknowns in overconfident domains compared to a 1.3% reduction in calibrated/underconfident domains, $b = 7.2$, 95% CI [-0.4, 14.8], $p = .063$. In the CTA condition, neither of the simple effects was significant and there was no significant interaction, all p -values $> .5$.

Discussion

As in Study 2, considering the unknowns reduced confidence and overconfidence, but did not affect percent correct. The robustness of these effects to 2AFC vs. 4AFC, within- vs. between-participants and with randomly vs. non-randomly sampled questions suggests that considering the unknowns is an effective debiasing technique under a variety of conditions. Importantly, considering the unknowns selectively reduced confidence in domains where

participants were overconfident. We found some evidence that considering the alternative has some efficacy at reducing overconfidence (consistent with Koriat, Lichtenstein, and Fischhoff 1980), but the effect of this manipulation was not consistent across our studies. We found some increase in percent correct in Study 2 but no effect on confidence and a small effect on confidence in Study 3 but no effect on percent correct. Across the two studies, considering the unknowns was more effective than considering the alternative at reducing confidence and equal to or better at reducing overconfidence.

GENERAL DISCUSSION

Our studies show that the evaluation of what evidence is unknown or missing is an important determinant of judged confidence. However, people tend to underappreciate what they don't know. Thus, overconfidence is driven in part by insufficient consideration of unknown evidence.

We conceptualize known unknowns as evidence relevant to a probability assessment that a judge is aware that he or she is missing while making the assessment. We distinguish this from unknown unknowns, evidence that a judge is not aware he or she is missing. It is useful at this point to further distinguish two varieties of unknown unknowns. In some cases a judge may be unaware that he or she is missing evidence but could potentially recognize that this evidence is missing if prompted. We refer to these as retrievable unknowns. In other cases, a judge is unaware that he or she is missing evidence and furthermore would need to be educated about the relevance of that evidence in order to recognize it as missing. We refer to these as unretrievable unknowns. To illustrate the importance of these distinctions, consider again the assessment of how likely it is that Iraq possesses nuclear weapons. In making this judgment, an intelligence

analyst may explicitly ask herself whether Iraq possesses enriched uranium. The analyst may recall that enriched uranium is an important requirement for nuclear weapons, and that this factor is unknown. In this case, the question of whether or not Iraq has enriched uranium would be a known unknown. Alternatively, it may be that the analyst understands the relevance of uranium enrichment but does not consider this factor when judging the possibility of nuclear weapons. In this case the presence of enriched uranium is a retrievable unknown. Studies 2 and 3 demonstrate the effectiveness of using a prompt to direct attention to retrievable unknowns that people may not otherwise consider, as a means of reducing misplaced confidence and improving calibration. However, consider further a non-expert who does not know that enriched uranium is an important ingredient in nuclear weapons. In this case the presence of enriched uranium is an unretrievable unknown that a “consider the unknowns” prompt could never elicit, though presumably the novice could be educated. This analysis predicts that a “consider the unknowns” prompt will only be effective in reducing misplaced confidence to the extent that the judge has sufficient expertise to recognize unknowns when prompted to do so.³

Our results suggest a potent new method that could be disseminated to practitioners for reducing overconfidence. First, ‘considering the unknowns’ could be a self-administered treatment before making important judgments in situations where overconfidence is prevalent, such as when a CEO is making an acquisition (Malmendier and Tate 2005), when a CFO is

³ Evidence that is recognized to be unknown may also vary in terms of its specificity. For example, when predicting the outcome of a football game a judge might consider that the health of the starting quarterback for the home team has been in question and so it is unknown whether or not the backup will have to carry the offense—a specific known unknown. Alternatively, a judge might consider that the variables that determine how the respective offenses and defenses of the teams match up is beyond his or her knowledge—a general known unknown. When debiasing using our “considering the unknowns” prompt, it is not clear to us which class of unknowns, the specific or the general, will tend to have a stronger effect on confidence and calibration.

budgeting for an upcoming year (Ben-David, Graham, and Harvey 2007), or when a head of state is considering a military action (Johnson 2004).

Considering the unknown may also be a more effective debiasing technique than considering the alternative in some situations. In Studies 2 and 3 we compared ‘consider the unknowns’ to ‘consider the alternative’ and found that considering the unknowns was more successful in reducing overconfidence. Further, we have provided some evidence that considering the unknowns selectively reduces confidence only when people are overconfident, whereas there is no evidence to suggest that correction is selective when considering the alternative. Considering the unknowns may also be more effective than considering the alternative in judgment tasks where no obvious alternative exists. For instance, when estimating quantities in confidence intervals, such as ‘the cost of an advertising campaign’ an instruction to “consider the alternative(s)” does not make sense (Alpert and Raiffa 1982). However, it may be possible to reduce interval overconfidence in such cases by prompting judges to consider the unknowns. This may be a fruitful area of future study because overconfidence is pervasive in confidence intervals estimation, with few techniques available to fully eliminate overconfidence biases (Klayman et al. 1999; Moore and Healy 2008; Soll and Klayman 2004).

Although we tout the potential of implementing a “consider the unknowns” strategy for debiasing, we do not claim that it will always outperform considering the alternative. One reason is that “consider the alternative” can sometimes not only lead to reductions in confidence but also improvements in the proportion of items answered correctly (as we saw in Study 2). Additionally, considering the alternative may be a more viable approach when trying to debias others since it may be more compelling to argue for a concrete alternative option (playing “devil’s advocate”) than to argue that the other person is missing information (given that another

person's retrievable unknowns are not necessarily retrievable to the persuader). Of course these strategies are not mutually exclusive, and a hybrid strategy of considering both the *unknowns* and the *alternative* may be more effective than either strategy alone.

Donald Rumsfeld, secretary of defense during the invasion of Iraq in 2003 is famous for distinguishing between known knowns, known unknowns, and unknown unknowns. Our research suggests that the administration's overconfidence that Saddam Hussein possessed weapons of mass destruction may have been due, in part, to focusing too much on the known knowns and neglecting the known unknowns. When Colin Powell made a speech to the UN Security Council in February of 2003 in which he presented a persuasive series of known facts supporting the existence of WMDs in Iraq he stated, "My colleagues, every statement I make today is backed up by sources, solid sources. These are not assertions. What we're giving you are facts and conclusions based on solid intelligence." If Colin Powell wanted his audience to have a more balanced view, he should have also articulated what was unknown to the Bush administration. Known unknowns could have ultimately strengthened or weakened the case for WMDs once they were resolved. For example, US officials might have explicitly acknowledged how little they understood about Mr. Hussein's possible motivations for remaining coy about his nuclear program, and moderated their confidence. Recently it has come to light that Mr. Hussein was far more concerned about an internal coup or a Shiite rebellion than he was about a U.S. invasion, and so he encouraged everyone—from opponents in Iran to his own generals—to believe that he might have WMDs (Gordon and Trainor 2006). Our studies suggest there would have been little downside to US officials considering what is unknown, at least from a judgment perspective. If unknowns are high, considering the unknown just might reduce overconfidence and if unknowns are low, considering unknown evidence will not impact calibration.

Chapter 2: Two Dimensions of Uncertainty Predict Investment Strategy

ABSTRACT

Estimating the future returns of any stock investment is an exercise fraught with uncertainty. This uncertainty can be attributed to failures in knowledge (e.g., being unfamiliar with an industry or company; what we call *epistemic uncertainty*) or to random processes (e.g., asset volatility; what we call *aleatory uncertainty*). In this paper we examine real investment behaviors and find that these two dimensions of subjective uncertainty are related to different investment decisions. To the extent that investors attribute stock market uncertainty to failures in knowledge they tend to seek financial advice and information. In contrast, to the extent that investors attribute stock market uncertainty to random processes they tend to seek to limit their exposure to any one asset through diversification. In investment decisions with real money at stake we show that knowledge predicts when investors will be tolerant of epistemic uncertainty whereas traditional measures of risk preference predict when investors will be tolerant of aleatory uncertainty. Perceptions of aleatory uncertainty can be primed to improve financial decisions, as reflected in lower portfolio concentration and less overconfident financial judgments, but only among the financially unsophisticated, providing an intervention that could improve financial decision making for those who need it most.

How to invest one's assets is among the most important financial decisions a person can make, and different people approach this decision in dramatically different ways. To illustrate, imagine two investors, Jack and Jill. Jack believes that the market rewards talent and expertise: with the right investment strategy and knowledge, he feels confident that one can consistently identify winning and losing assets and outperform the market. As a result, Jack spends considerable amounts of time and money researching and trading individual assets. In contrast, Jill views the market as profoundly stochastic, and places little confidence in her (or anyone else's) ability to reliably pick winners and losers. Instead, Jill focuses her energy on maintaining many different and diversified assets that tend to follow aggregate market trends over long periods of time, which she views as reducing noise (and therefore her exposure to risk) in a fundamentally random environment.

In this paper we suggest that differences in investment strategy can arise from differences in how people view the nature of market uncertainty. The distinction between epistemic and aleatory conceptions of uncertainty has longstanding roots the development of probability theories (Hacking 1975), but has only recently been empirically investigated as a behavioral phenomenon (Tannenbaum, Fox, and Ülkümen 2016; Ülkümen, Fox, and Malle 2016). In this paper, we examine the association between perceptions of epistemic and aleatory uncertainty in the stock market and investment behaviors, beliefs, and strategies.

While behavioral models usually treat uncertainty as a unitary construct, in recent decades theorists have occasionally proposed psychological variants of uncertainty (e.g., Fox and Ülkümen 2011; Howell and Burnett 1978; Kahneman and Tversky 1982; Teigen 1994). In particular, Fox and Ülkümen (2011) distinguish uncertainty attributed to deficiencies in one's knowledge, information, or skills in assessing an event that is, in principle, knowable (*epistemic*

uncertainty) from uncertainty attributed to the behavior of fundamentally random or chance processes (*aleatory* uncertainty), with epistemic and aleatory uncertainty constituting independent dimensions.

Recent empirical studies have validated the intuitive distinction between epistemic and aleatory dimensions of uncertainty. In particular, research has found that forecasters are assigned more credit for correct predictions and more blame for incorrect predictions to the extent that relevant uncertainty is seen as epistemic, whereas they are deemed more lucky (unlucky) for correct (incorrect) predictions to the extent that uncertainty is seen as aleatory (Fox, Tannenbaum, Ülkümen, Walters & Erner, in preparation). People tend to communicate degrees of epistemic (knowable) uncertainty using expressions such as “90% sure” or “fairly confident” whereas they tend to communicate degrees of aleatory (random) uncertainty using expressions such as “90% chance” or “high probability” (Ülkümen, Fox, and Malle 2016). Forecasters tend to make more extreme probability judgments, holding evidence strength constant, the more epistemic they perceive relevant uncertainty to be, and they tend to make more regressive judgments the more aleatory they perceive uncertainty to be (Tannenbaum, Fox, and Ülkümen 2016). Aversion to ambiguity (Ellsberg 1961) appears to be driven by reluctance to bet in situations where the decision maker feels relatively ignorant, especially to the extent that relevant uncertainty is seen as more epistemic and less aleatory (Fox, Goedde-Menke & Tannenbaum, in progress).

In the context of investing, it is generally accepted that investors seek to maximize expected returns while minimizing uncertainty in possible returns (e.g., Markowitz 1952). We propose that the uncertainty around any given financial outcome can be perceived as epistemic, aleatory, or both. Importantly, we suggest that investor perceptions of epistemic versus aleatory

uncertainty in stock market outcomes suggest distinct corrective measures. Because epistemic uncertainty is generally attributed to missing knowledge or skill, we expect that investors will attempt to remedy epistemic uncertainty by seeking information or consulting experts. Moreover, we expect that investors who see the market as more epistemic will be more sensitive to changes in their level of knowledge when making investment decisions.

In contrast, because aleatory uncertainty is attributed to random and inherently unpredictable sources, additional information will not be helpful in reducing it. Instead, we expect that investors will attempt to address aleatory uncertainty by reducing asset concentration and spreading risk across different (imperfectly correlated) investments. Moreover, we expect that investors who see the market as more aleatory will be more affected by their general risk preference when investing. We predict, further, that greater perceptions of aleatory uncertainty will be associated with wider ranging price estimates as investors perceive these assets to be more volatile.

In this paper we present five studies that are designed to demonstrate the association between perceived epistemic and aleatory dimensions of market uncertainty and investor behaviors. In Study 1 we examine real investment decisions in a panel of real investor to show that investors that perceive stock market uncertainty to be more epistemic seek financial advice whereas investors that perceive stock market uncertainty to be more aleatory seek diversification. We next demonstrate these same effects in an experimental context with real money at stake by showing we can push real investors to reduce asset concentration and by priming aleatory uncertainty in Study 2, seek more information by priming epistemic uncertainty in Study 3, and invest more in assets believed to be governed by epistemic uncertainty by increasing information in Study 4. Finally, we provide financial advisors a roadmap for tailoring investment guidance

across these dimensions by showing that appetite for epistemic uncertainty is moderated by knowledge in Study 4 whereas appetite for aleatory uncertainty is moderated by risk preference in Study 5.

STUDY 1

In our first study we examine decisions people have made in their own investment portfolios. We recruited a sample of investors and assessed perceptions of stock market uncertainty along with portfolio composition and investment behaviors. We predicted that perceptions of greater epistemic uncertainty in the stock market would be associated with placing higher value on stock advice whereas perceptions of greater aleatory uncertainty would be associated with lower stock concentration.

Method

We recruited participants from a large, diverse Qualtrics panel to complete a survey in exchange for \$10. The Qualtrics panel is comprised of over 525,000 members ranging in age from 18 to 50 with a broad range of professional experience. Before completing the questionnaire, participants were screened for adequate financial experience. To be eligible for the study, participants were required to own more than \$1,000 in stock market investments, be between the ages of 18 and 65, report making their own investment decisions, rate their knowledge of the stock market as a three or higher on a five point scale (3 = *I know what investing in the stock market is but do not consider myself very knowledgeable in the subject*, 4 = *I know what investing in the stock market is and have a moderate level of knowledge in the subject*, 5 = *I consider myself an expert on stock market investing*). To be included in the study we also required participants to correctly answer two simple financial screening questions. A

complete list of the questions and required responses can be found in Appendix B. Of the 7,191 individuals who responded to the initial screening questions, 354 passed the initial screening procedure.

EARS rating. Participants evaluated stock market uncertainty using a 6-item version of the Epistemic-Aleatory Rating Scale (EARS) that we have developed and validated elsewhere (Fox et al., 2017). Participants rated the uncertainty associated with the “approximate return of an individual stock over 1 year.” The scale prompted participants to rate their agreement on 7-point scales (1 = *strongly disagree*, 7 = *strongly agree*) with a set of statements that measured both feelings of epistemic uncertainty (e.g., “the approximate return of an individual stock over 1 year is something that becomes more predictable with additional knowledge or skills”) and aleatory uncertainty (e.g., “the approximate return of an individual stock over 1 day is something that has an element of randomness”). This scale can be viewed in Appendix B.

Financial Advisor. Participants reported whether they currently did or did not have a financial advisor. If a participant did have a financial advisor, they then reported the fee they pay to the financial advisor as a percentage of assets under management.

Concentration. Participants next reported the number individual stocks they currently held. We define holding a single stock as a more concentrated stock portfolio and holding more stocks as a less concentrated stock portfolio. We excluded investors that do not hold any stocks and winsorize the data at a maximum of 100 stocks (meaning that values of more than 100 stocks were transformed to 100 stocks). Of our 354 respondents, 19 reported holding no individual stocks and 5 reported holding more than 100 individual stocks.

Trading frequency. Participants next reported the total value of their investments, the number of changes they make to their investments on a 7-point scale from “more than every day”

(1) to “fewer than one change every 12 months” (7), and the average period of time that they hold stocks and mutual funds on a 6-point scale from “several hours” (1) to “several days” (6).

Risk perception. As a control variable, we measured risk perception using 3 items from the risk perception scale of the DOSPERT (Weber, Blais, and Betz 2002; Appendix B).

Other measures. Participants also reported the percentage of their investment assets they hold in each of the following categories: individual stocks, stock mutual funds, stock index funds, individual bonds, bond mutual funds, bond index funds, individual commodities, commodities mutual funds, commodities index funds, individual real estate, real estate mutual funds, real estate index funds. They then reported the total value of these investment assets on a 1-7 scale. Participants next reported the percentage of non-investment assets they held in their home, pension, annuities, cash, and other (open response). They then reported the total value of these other assets on a 1-7 scale. Participants then completed the simple 3-item financial literacy test (Lusardi, Mitchell, and Curto 2010). Finally, participants provided demographic information and were debriefed.

Results

EARS rating. As expected, participants perceived two independent dimensions of market uncertainty. The epistemic items exhibited high reliability (Cronbach’s $\alpha = .801$), as did the aleatory items (Cronbach’s $\alpha = .796$). We further examined the psychometric properties of the EARS with a confirmatory factor analysis with oblique rotation. Confirmatory factor analysis suggested that a two-factor model fit the EARS well, with an eigen value of 2.56 for factor 1, .81 for factor 2, and a drop off at factor 3 to an eigen value of .06. Only the epistemic items loaded on factor 1 (epistemic 1, .818; epistemic 2, .707; epistemic 3, .612; aleatory 1, .241; aleatory 2, .071; aleatory 3, -.151) whereas only the aleatory items loaded on factor 2 (epistemic 1, -.013;

epistemic 2, .047; epistemic 3, .006; aleatory 1, .579; aleatory 2, .502; aleatory 3, .942). These two factors were correlated, $r = .487$.

Financial Advisor. We first examined the relationship between the perception of uncertainty on the EARS and expertise. We predicted that only perceptions of epistemic uncertainty would be correlated with reliance on financial advice, with greater epistemic uncertainty being associated with more reliance on expert advice. To examine this, we regressed whether or not a respondent paid a financial advisor for advice (1 = yes; 0 = no) on perceptions of epistemic uncertainty, aleatory uncertainty while controlling for risk perception, percentage of assets held in stocks, financial literacy, total investment asset value, other assets value, and number of stocks held using a robust logistic regression. As expected, participants who perceived epistemic uncertainty were more likely to have financial advisors, $b = .30$, 95% CI = [.09, .50], $p < .01$. In contrast, the perception of aleatory uncertainty, $b = -.02$, 95% CI = [-.26, .23], $p = .985$ did not reliably predict having a financial advisor. As predicted, a test of the absolute difference between the epistemic coefficient and aleatory coefficient reveal the epistemic coefficient is significantly larger, $X^2 = 4.53$, $p = .033$.

Concentration. We next examined whether investors perceiving more aleatory uncertainty in the stock would hold less concentrated stock portfolios (presumably to spread risk across multiple assets). To do so, we regressed portfolio stock concentrations on epistemic uncertainty, aleatory uncertainty, while controlling for risk perception, percentage of assets held in stocks, financial literacy, total investment asset value, other assets value, and whether the participant had a financial advisor. As predicted, participants who perceived greater aleatory uncertainty in the market held a greater number of stocks (i.e., less concentrated stock portfolios), $b = 3.28$, 95% CI = [1.42, 5.14], $p = .001$. In contrast, perceptions of epistemic

uncertainty did not reliably predict portfolio stock concentration, $b = .83$, 95% CI = [-.79, 2.44], $p = .316$. As predicted, a test of the absolute difference between the epistemic coefficient and aleatory coefficient reveal the aleatory coefficient is significantly larger, $X^2 = 11.09$, $p = .001$.

Financial literacy. Finally, we examined a regression with financial literacy as the dependent variable and epistemic and aleatory uncertainty as the independent variables. We found that participants that were higher in financial literacy rated the stock market as less epistemic, $b = -.24$, 95% CI = [-.32, -.16], $p < .001$, and less aleatory, $r = -.11$, $p < .05$.

Discussion

The results of Study 1 suggest that perceptions of epistemic and aleatory dimensions of uncertainty have distinct consequences for investor behavior. Confirming our predictions, investors who perceive stock market uncertainty to be more epistemic were more likely to seek and pay for financial advice. In contrast, investors who perceive stock market uncertainty to be more aleatory held less concentrated portfolios. Apparently, investors seek to address epistemic uncertainty by consulting experts and address aleatory uncertainty through diversification.

We also find that greater perceptions of both epistemic and aleatory uncertainty were associated with lower financial literacy. To further examine this result we compared perceptions of epistemic uncertainty in this study to those in a study of 37 practicing financial advisors (recruited through an executive education program at UCLA) who also evaluated stock market investing using a four-item version of the EARS. In a separate sample t-test we find that financial advisors perceived lower epistemic uncertainty in stock investing ($M = 2.90$), compared to non-professional investors ($M = 3.99$; $t = 4.49$, $p < .001$). However, perceptions of aleatory uncertainty did not reliably differ between financial advisors ($M = 4.47$) and non-professional investors ($M = 4.70$; $t = 1.09$, $p = .28$). This result, taken together with our finding that investors

lower in financial literacy perceive the stock market to be more epistemic, indicates that less experienced investors overestimate the predictability of the stock market.

STUDY 2

We found in Study 1 that perceptions of epistemicness and aleatoriness in stock movement predicted investment behaviors. In Study 2 we extend this finding in an experimental design to investigate whether perceptions of uncertainty can be manipulated and if they can thereby perturb investment behaviors. We sought to experimentally prime perceptions of epistemic (knowable) uncertainty by asking participants to elaborate their forecast of company performance using a single scenario or prime perceptions of aleatory (random) uncertainty by asking participants to elaborate their forecast of company performance using a set of possible alternative outcomes. Prompting individuals to engage in such “singular” versus “distributional” reasoning has been found to influence perceptions of epistemic and aleatory uncertainty (Tannenbaum, Fox, and Ülkümen 2016). We expected lower portfolio concentration and wider confidence intervals around forecasted earnings when primed with distributional reasoning compared to singular reasoning.

Method

We recruited participants from a large, diverse Qualtrics panel to complete a brief survey in exchange for \$8. As in Study 1, participants were screened for adequate financial experience. To be eligible for the study, participants were required to own more than \$1,000 in stock market investments, be between the ages of 18 and 65, report making their own investment decisions, rate their knowledge of the stock market as a three or higher on a five point scale (3 = *I know what investing in the stock market is but do not consider myself very knowledgeable in the*

subject, 4 = *I know what investing in the stock market is and have a moderate level of knowledge in the subject*, 5 = *I consider myself an expert on stock market investing*). To be included in the study we also required participants to correctly answer two simple financial screening questions. A complete list of the questions and required responses can be found in Appendix B. Of the 1,551 individuals who responded to the initial screening questions, 201 qualified for participation.

Prime. All participants read the 2014 Q3 earnings release for two companies, Amazon.com and Ford, in a randomized order (Appendix B). We used Ford and Amazon because we assumed most participants would be familiar with these companies, and because their press releases were relatively non-technical and easy to understand.

We randomly assigned participants to one of two experimental conditions. To prompt *singular* reasoning, we asked one group of participants to construct a single scenario (12-15 sentences long) for each company about how much revenue would be generated in Q3 2015. To prompt *distributional* reasoning, we asked a second group of participants to construct three possible scenarios (each 4-5 sentences long) for each company about how much revenue would be generated in Q3 2015. We then asked all participants to make 90% confidence interval estimates for revenue of these two companies.

In order to ensure thoughtful responding, we removed 29 participants who failed to adequately engage in the writing task. We would expect an average sentence to have 15-20 words and 6-7 characters per word including spaces (Cutts 2013). Thus, we would expect 12-15 sentences to equate to minimum of 1,080 characters, on average. We took 400 characters as a conservative minimum and eliminated any participant with descriptions of less than 400 characters averaged across Amazon.com and Ford.

EARS. Participants completed the writing task for the first randomly selected company and on the same page participants assessed the underlying uncertainty associated with the forecasting task on the 6-item version of the EARS, responding to the prompt: “Consider the task of evaluating the approximate next year sales amount of [COMPANY]. The approximate next year sales amount of [COMPANY]...” (Appendix B). Participants also forecasted the company’s revenue for the following year, and were also asked to place 90% confidence intervals around their estimate. As a manipulation check, Participants then completed the same writing task, revenue estimate, and EARS assessment for the second company.

Concentration. Next, we asked participants to imagine investing \$100 dollars between the two companies. As an incentive to respond thoughtfully, we informed participants that some respondents would be randomly selected to receive the actual payout of their investment decisions after six months.

Financial literacy. Finally, participants completed the sophisticated financial literacy test (Lusardi, Mitchell, and Curto 2010) and were debriefed.

Results

Manipulation check. To verify that our intervention was successful we examined the impact of experimental condition on epistemic and aleatory rating. As predicted, participants primed with distributional reasoning reported lower epistemic ratings than participants primed with singular reasoning ($M_s = 4.34$, $CI = [4.14; 4.54]$ vs 4.65 , $CI = [4.45; 4.85]$; $t(170) = 2.21$, $p = .028$). Participants in the distributional prime condition also reported greater aleatory market uncertainty than participants in the singular prime condition, although this difference failed to reach statistical significance ($M_s = 4.79$, $CI = [4.56; 5.02]$ vs 4.62 , $CI = [4.40; 4.84]$; $t(170) = 1.05$, $p = .295$).

Revenue estimates. We next examined revenue change estimates. We predicted that the distributional prime would increase aleatory thinking (relative to epistemic thinking) and expand confidence interval widths, whereas the singular prime would increase epistemic thinking (relative to aleatory thinking). Confidence interval widths were standardized for each participant by taking the difference between the upper and lower bounds and dividing by expected earnings. This was done separately for each company, and for each participant we took the average of the two standardized interval widths. Confirming our prediction, participants prompted to think distributionally provided wider confidence intervals than participants prompted to think singularly ($M_s = 0.24$, $CI = [.19, .28]$ vs 0.18 , $CI = [.15, .20]$; $t(170) = 2.60$, $p = .031$).

Concentration. We next examined portfolio concentration, which was calculated by taking the percentage of money placed in the largest investment. Confirming our prediction, participants primed with a distributional mindset concentrated their investments less than participants primed with a singular mindset ($M_s = 75.6\%$ vs 80.3% ; $t(170) = 2.15$, $p = .033$).

Next, we examined whether our findings interacted with participants' degree of financial literacy, which we operationalized as the percent correct across our 11 financial literacy questions. Using a fractional response model (Papke and Wooldridge 1993), we regressed portfolio concentration on experimental prime (0 = distributional, 1 = singular), financial literacy scores, and the interaction term between the two predictors. As illustrated in Figure 4, our singular vs. distributional mindset prime had the strongest effect on participants low in financial literacy ($b = -75.6$, $CI = [-125.9, -25.3]$, $p = .003$ for the interaction term). Based on the average marginal effects from our model, participants relatively low in financial literacy (1 *SD* below the mean) would be expected to show a 24 percentage-point decrease in portfolio concentration when primed to think distributionally rather than singularly (predicted probabilities were 66.7%

vs 42.5%; $b = 0.24$, $SE = 0.5$, $p < .001$). In contrast, participants relatively high in financial literacy (1 SD above the mean) would not be expected to show a reliable difference in their portfolio behavior across the two priming conditions (predicted probabilities were 60.0% vs 54.4%; $b = -.06$, $SE = 0.06$, $p = .36$). Financial literacy scores did not interact with the experimental prime on any of the other dependent variables.

Discussion

Study 2 demonstrates that perceptions of market uncertainty can be reliably perturbed by prompting individuals to consider a single scenario versus multiple possible scenarios. Participants were less likely to diversify their portfolios when prompted to think about singular cases, which reliably influenced perceptions market uncertainty. We also find that our experimental prime had a larger influence on those with the lowest levels of financial literacy — presumably those who were least informed about how to appropriately allocate their investments. Thus, shifting beliefs about epistemic and aleatory market uncertainty may prove to be a useful intervention for improving financial decision making among those who need it the most. One shortcoming of this study is that we did not find that aleatory ratings mediated the relationship between condition and diversification. This may have occurred because our prime was not strong enough to be picked up by our EARS measurement tool. We address this in Study 3 where we do find partial mediation.

STUDY 3

We showed in Study 1 that perceptions of epistemic and aleatory market uncertainty were related to distinct investment behaviors. In Study 2 we showed that priming a more aleatory mindset leads people to greater diversification of their investments. In Study 3 we build on these

findings to show that we can push investors to seek financial information by increasing perceptions of epistemic uncertainty. We sought to experimentally manipulate perceptions of stock market uncertainty by asking participants to read a short article that focused people on the importance of knowledge in investing (which we expected to enhance perceptions of epistemic uncertainty among respondents) or how experts are often incorrect (which we expected to diminish perceptions of epistemic uncertainty among respondents). We predicted that participants primed with greater epistemic uncertainty would be more willing to pay for information relevant to a stock investment.

Method

We recruited participants on Amazon Mechanical Turk ($N = 397$) for \$.50 each plus a performance incentive. We eliminated 47 participants (11.8%) because they did not pass an attention check. Participants were asked to complete a short investment task and told that some participants would be selected to receive the value of their investment 6 months in the future.

Prime. Participants were randomly assigned to one of two conditions. Participants read either an article about Peter Lynch, a successful mutual fund manager (high epistemicness prime) or an article about the inaccuracy of financial gurus (low epistemicness prime). Both can be found in Appendix B.

EARS. Participants first assessed uncertainty associated with the return of an individual stock on the 6-item version of the EARS with the prompt: “Consider the task of evaluating the return of a typical stock over the course of the next 6 months. The return of a stock is....”

Next, participants learned about the investment task that they would be completing, reading the statement:

“You will be investing \$50 over the next 6 months. You may invest this money across three stocks: Aon PLC, Nabors Industries, and Textron Inc. You will first be given the opportunity to read a research report on these companies. Some participants will be randomly chosen to have these investment decisions honored for real money.”

Information Search. Participants were next given the opportunity to buy research reports on all three companies for \$2, which would be deducted from the value of their investments at the end of 6 months if they were randomly selected to be a “real money” participant. Participants who chose to buy the report then viewed a real research report that combined analyst estimates on the price targets of these companies over the following 6 months. The report was constructed by the researchers based on real data obtained from IBIS (see Appendix B).

Participants next made an investment allocation across Aon PLC, Nabors Industries, and Textron Inc. Finally, participants reported their investment experience on a 5-point rating scale, provided demographic information, and were debriefed.

Results

Manipulation check. To verify that our epistemicness prime was effective we examined the impact of the manipulation on epistemic and aleatory ratings. As predicted, epistemicness ratings were higher after participants read the article designed to prime perceptions of high epistemicness, $M_s = 4.5$, 95% CI = [4.4, 4.7] than after participants read the article designed to prime perceptions of low epistemicness, $M_s = 3.6$, 95% CI = [3.4, 3.6], $t(318) = 7.08$, $p < .001$. In addition, perceptions of aleatoriness were higher after participants read the article to prime low epistemic perceptions, $M_s = 5.4$, 95% CI = [5.2, 5.6] than after the article to prime high epistemic perceptions, $M_s = 4.9$, 95% CI = [4.7, 4.0], $t(318) = 4.53$, $p < .001$. To examine whether the shift in perceptions of aleatoriness was smaller than the shift in perceptions of

epistemicness we conducted a repeated measure regression with the [aleatory/epistemic] rating as the dependent variable and the condition, rating type (whether the rating was epistemic or aleatory) and the interaction of condition and rating type as independent variables while clustering by participant. Aleatory rating was reverse coded since the high epistemic prime tended to suppress perception of aleatoriness. As predicted, we found a significant interaction, $b = -.39$, 95% CI [-.70, -.09], $p = .01$, indicating that the shift in epistemic rating across conditions was greater than the shift in aleatory rating.

Information Search. We next examined information search. We predicted that participants primed to think stock investing is more epistemic would search for more information. Confirming our prediction, a smaller percentage of participants searched for information after reading the article designed to prime low epistemicness, $M_s = 43.2\%$, 95% CI = [35.3%, 51.1%], compared to the article to prime high epistemicness $M_s = 73\%$, 95% CI = [66.3%, 80.0%], $p < .001$. We next examined the relationship between information search and the perception of uncertainty in a logistic regression with information search as the dependent variable and with epistemicness and aleatoriness ratings as independent variables. As predicted, we found that a higher epistemicness ratings were related to more information search, $b = .34$, 95% CI = [.15, .54], $p < .001$, whereas aleatoriness ratings were not significantly related, $b = -.01$, 95% CI = [-.23, .21], $p = .933$.

Mediation analysis. We next examined the extent to which perceptions of epistemicness mediated information search behavior. To do this we conducted a bootstrap mediation analysis for dichotomous outcomes, in which epistemicness rating was the mediator between condition and information search while controlling for aleatoriness rating. First, we found that experimental condition was a reliable predictor of information search, $b = 1.28$, 95% CI = [.80,

1.76], $p < .001$. Second, we found that condition was a reliable predictor of epistemicness rating, $b = .79$, $p < .001$, and epistemicness rating predicted information search, $b = .34$, 95% CI = [.15, .54], $p < .001$. Third, when both condition and epistemic rating were included in the model, we found that epistemicness rating predicted information search, $b = .21$, 95% CI = [.00, .41], $p < .05$, and experimental condition predicted information search, $b = 1.14$, 95% CI = [.63, 1.64], $p < .001$. Overall, we found that epistemic rating explained 12.6% of the experimental treatment effect on information search, $b_{\text{indirect}} = .04$, 95% CI = [.00, .08] (bootstrapping based on 500 resamples), indicating partial mediation.

We also examined a bootstrap mediation analysis for dichotomous outcomes, in which aleatoriness rating was the mediator between condition and information search while controlling for epistemic rating. We found that aleatory rating explained 0.0% of the experimental treatment effect on information search, $b_{\text{indirect}} = -.01$, 95% bootstrapped 95% CI = [-.03, .02] (bootstrapping based on 500 resamples), indicating no mediation.

Discussion

In Study 3 we found that perceptions of market uncertainty can be perturbed by prompting individuals to read a short article. Participants were more likely to search for information after reading the article that primed a greater perception of epistemicness. We find that perception of epistemic uncertainty, as measured on the EARS, partially mediates the relationship between condition and information search. In contrast, the perception of aleatory uncertainty was not related to information search. This provides some support for the notion that the perception of more epistemic uncertainty in stock prices drives investors to seek more information whereas the perception more or less aleatory uncertainty does not.

STUDY 4

In Study 1 we showed that willingness to pay for financial advice is associated with attributions of greater epistemic uncertainty. In Study 4 we extend these findings in an experimental setting by manipulating financial advice, using a familiar stock (Apple). We predict that financial advice will be more influential on investment choices when an investor perceives uncertainty in a stock's returns as more epistemic in nature.

Method

We recruited participants using Amazon's Mechanical Turk ($N = 302$) for \$.50 each plus a performance incentive. We eliminated 76 participants (25.2%) because they did not pass an attention check.

First round judgments and decisions. To establish a baseline of how informed participants considered themselves concerning future movement of Apple stock, we asked: "Please rate how much information you feel you have about how the stock price of Apple will change over the next 6 months" on a scale from 1 (No information) to 7 (A lot of information). Participants next indicated whether they thought the stock price of Apple in 6 months time would be: (1) above its current price of \$140 or (2) the same or below as its current price of \$140. Participants then chose between receiving either: (a) \$50 for sure, or (b) \$150 if their prediction about the future movement of Apple stock was correct (and \$0 otherwise). Participants were informed that some participants would be randomly selected to receive the outcome of their choice for real money.

Advice Manipulation. In the second phase of the experiment, participants were presented with a real 1 page research report that predicted the price of Apple stock would increase in the coming months. This report can be viewed in Appendix B.

Second round judgments and decisions. Participants completed all of the same measures in the same order and format as round 1, while viewing the analyst stock report.

EARS. Participants completed the 6-item EARS in connection with “the stock price of Apple over the next 6 months” and were then debriefed.

Results

We examined our prediction that the financial advice would have a greater influence across our three repeated measures: knowledge, prediction and choice.

Information. We predicted that participant who viewed the uncertainty associated with future movement of Apple stock as more epistemic would show a larger increase in self-rated stock information after viewing the analyst report. We conducted this within subject analysis in a regression with rated information as the dependent variable and round (time 1 vs. time 2), epistemic rating, aleatory rating, the interaction of round and epistemic rating, and the interaction of round and aleatory rating as the independent variables. Standard errors were clustered by participants.

We first examine the marginal effect of round 1 vs. 2 and found that mean information rating increased from 2.74 in the first round to 4.76 in the second round, $b = 2.01$, 95% CI = [1.78, 2.23], $p < .001$, indicating that our information manipulation increased participants’ perceptions of their information about Apple. We next examined the interaction between round (1 vs 2) and epistemicness rating. As predicted, we found a significant interaction, $b = .45$, 95% CI = [.25, .64], $p < .001$, confirming our prediction that participants who viewed Apple stock as more epistemic would exhibit a larger increase in ratings of subjective information between rounds 1 and 2. In contrast, the interaction between round (1 vs. 2) and aleatoriness rating was marginally significant $b = .19$, 95% CI = [-.02, .41], $p = .082$. A test of the difference between

coefficients reveals that the interaction between round and epistemicness ratings is larger than the interaction between round and aleatoriness ratings, $X^2 = 3.82$, $p = .052$.

Prediction. We hypothesized that participants who viewed the uncertainty associated with Apple stock as more epistemic would be more likely to change their prediction to accord with the stock analyst's (bullish) recommendation. We conducted this analysis as a logistic regression with stock prediction as the dependent variable and round (1 vs. 2), epistemicness rating, aleatoriness rating, the interaction of round and epistemicness rating, and the interaction of round and aleatoriness rating as the independent variables. As predicted, we found an interaction between round and epistemicness rating, $b = -.30$, 95% CI = [-.61; .01], $p = .061$, indicating that participants who viewed the uncertainty concerning future Apple stock price as more epistemic were more likely to change their prediction to be consistent with the stock analyst's recommendation after reading the analyst's report. In contrast, there was no interaction between round and aleatoriness rating, $b = .00$, 95% CI = [-.32; -.33], $p = .986$. A test of the difference between coefficients reveals that the interaction between round and epistemicness ratings is larger than the interaction between round and aleatoriness ratings, $X^2 = 3.62$, $p = .057$.

Choice. We predicted that participant who perceived the uncertainty associated with Apple stock to be more epistemic would be more likely to act on the analyst's information in their investing behavior. We expected that participants who in the first round thought Apple would increase in price and maintained this belief in the second round ($n = 161$) would be more likely to bet on this prediction after reading the analyst report if they believed the uncertainty surrounding Apple Stock was more epistemic. Conversely, we predicted that participants who thought in the first round that Apple would decrease in price and maintained this prediction in the second round ($n = 21$) would be less likely to bet on this prediction after reading the analyst

report if they believed the uncertainty surrounding Apple Stock was more epistemic. We do not have a clear prediction of how the analyst report would impact the betting of participants who predicted in the first round that the price of Apple stock would decrease then changed their view in the second round ($n = 40$), or thought in the first round that the stock would increase, then changed their view in the second round ($n = 4$), so we excluded these participants from the analysis that follows. We next reverse coded the choice among participants who thought Apple would decrease in price in round 1 and maintained this belief in round 2, since we expected these participants to be less likely to choose the prospect after reading the analyst's report.

We then conducted this analysis using a logistic regression with choice of the Apple stock prospect (rather than the sure amount) as the dependent variable and round (1 vs. 2), epistemicness rating, aleatoriness rating, the interaction of round and epistemicness rating, and the interaction of round and aleatoriness rating as independent variables. As predicted, we find an interaction between round and epistemicness rating, $b = .19$, $CI = [-.01; .39]$, $p = .062$, indicating that participants who viewed the uncertainty of Apple stock as more epistemic were more likely to choose the Apple prospect after reading the analyst's report if they held the same prediction as the analyst (or less likely to choose the Apple prospect after reading the analyst's report if they maintained the opposing prediction of the analyst). In contrast, there was no interaction between round and aleatoriness rating, $b = -.03$, $95\% CI = [-.22; .15]$, $p = .708$. A test of the difference between coefficients reveals that the interaction between round and epistemicness ratings is larger than the interaction between round and aleatoriness ratings, $X^2 = 5.33$, $p = .021$.

Discussion

The results of Study 4 accord with our prediction that stock advice has a greater influence on investors who view stock price uncertainty as more epistemic, but not on those who view stock price uncertainty as more or less aleatory. Stated another way, we found that perception of epistemicness moderates participant's sensitivity to their level of information. In Study 5 we examine whether there is an analogous impact of perceived aleatoriness on sensitivity to risk preference.

STUDY 5

In Study 5 we sought to show that tolerance for aleatory uncertainty depends on tolerance for risk whereas tolerance for epistemic uncertainty does not. To the extent that movement in a stock is seen as random or unpredictable than people's preference to bet on predictions of stock movement should more closely resemble their preferences to bet on chance outcomes.

Method

We recruited participants on Amazon's Mechanical Turk ($n = 288$) for \$.50 each plus a performance incentive.

Risk Preference. We first assessed risk preference by asking participants to complete a short task in which they accepted or rejected two 50/50 prospects. Participants were told "Below you will find a choice between a sure gain and a 50/50 coin flip prospect. Please indicate if you prefer the sure gain or the coin flip prospect in the following scenario." Participants chose between "Gain \$50 for sure" or "If the coin turns up heads you gain \$150, if the coin turns up tails, then you gain \$0." If a participant selected the risky option, then the amount that could be gained if the coin lands heads was reduced to \$100 and the participant chose again. If the participant instead selected the safe option, the amount that could be gained if the coin lands

heads was changed to \$200 and the participants chose again. Thus, participants were titrated into 4 levels of risk aversion: Participants who accepted both prospects (risk seeking), participants who accepted the first prospect and rejected the second prospect (mildly risk averse) participants who rejected the first prospect and accepted the second prospect (moderately risk averse), and participants who rejected both prospects (strongly risk averse).

EARS. We next presented participants with eight stocks and asked them to assess the return of each stock over the coming week. Participants first read a 1 paragraph description of the companies taken from Reuters, then assessed the uncertainty associated with each of these stocks, one at a time, using the 6-item EARS.

Stock Prediction and Choice. Participants next predicted whether each of the stocks that they had previously evaluated would increase in value more or less than the S&P 500 over the following week. For each stock, participants then chose between the prospect of receiving either (a) \$30 for sure, or (b) \$90 if their prediction about the stock is correct (and \$0 otherwise). Participant were told that some respondents would be randomly selected to receive the outcome of one of their choices for real money. Finally, participants assessed the probability that their prediction about each stock would be correct, rated their knowledge of each stock, and then were debriefed.

Results

We examined our prediction that risk tolerance would be more diagnostic of choices concerning stocks whose returns were viewed as more aleatory. We ran a logistic regression clustered by participant with choice of the uncertain investment prospect (rather than the sure amount) as the dependent variable, and aleatoriness ratings, risk preference, epistemicness ratings, probability, the interaction between aleatoriness ratings and risk preference and the

interaction between epistemicness ratings and risk preference as the independent variables. As predicted, we found that willingness to accept the investment prospect significantly increases with the interaction between rated aleatoriness and risk tolerance, $b = .14$, 95% CI = [.01, .27], $p < .05$ whereas there is no significant interaction between rated epistemicness and risk tolerance, $b = -.08$, 95% CI = [-.20, .04], $p = .17$. A test of the difference between these coefficients reveals that the interaction between aleatoriness ratings and risk tolerance is significantly different from the interaction between epistemicness ratings and risk tolerance, $X^2 = 5.58$, $p < .05$.

Discussion

This study suggests that when a person believes a stock price is governed by more aleatory uncertainty, her willingness to bet on its future movement is more sensitive to her risk preferences.

GENERAL DISCUSSION

In this paper, we find that investor behaviors are influenced by their subjective perception of two independent dimensions of uncertainty. Investors who believe that stock returns are more epistemic tend to seek information and investment expertise to minimize this uncertainty whereas when investors who believe that stock returns are more aleatory tend to diversify their assets to minimize this uncertainty. In a panel of real investors, we found that perceptions of greater epistemicness were associated with a greater tendency to seek and value financial advice, whereas perceptions of greater aleatoriness were not. In contrast, perceptions of greater aleatoriness were associated with lower stock concentration whereas perceptions of epistemicness were not. We then demonstrated similar effects in an experimental context, with real money at stake. In Study 2 we primed perceptions of aleatoriness which led investors to

reduce their asset concentrations, and in Study 3 we primed perceptions of epistemicness which led investors to pay more for relevant information. In Study 4 we showed that investors were sensitive to their level of information to the extent they viewed relevant uncertainty as more epistemic. Finally, in Study 5 we showed that investors were more sensitive to their risk preference to the extent that they viewed relevant uncertainty as more aleatory.

While we are agnostic in this paper concerning the appropriateness of attributing stock market uncertainty to epistemic or aleatory factors, we suspect that people may have a tendency to perceive greater epistemic uncertainty in the market than is warranted. We note that experienced financial advisors attribute less epistemic uncertainty to the stock market than non-professional investors, and this hypothesis also accords with ample research in the judgment and decision making literature in which people tend to see more patterns where none exist (e.g., Gilovich, Vallone, and Tversky 1985; Tversky and Kahneman 1971). Perhaps more important, we speculate that perturbing perceptions of the nature of market uncertainty may have potential to improve investment decisions. Past research suggests that paying a financial advisor is a relatively bad investment strategy (e.g., Sharpe 1991) whereas diversification is the cornerstone of modern portfolio theory (Markowitz 1952). In addition, the efficient-market hypothesis (Basu 1977), a viewpoint held by many prominent professionals and academics in finance suggests that all publicly available information useful for predicting a future stock price has already been incorporated into the current stock price. Thus, there should be little epistemic uncertainty remaining for a stock advisor to resolve. Encouraging diversification based investment strategies and discouraging active management based strategies is likely to help most investors make better decisions. In Study 2 we showed that priming investors to think the stock market is more aleatory increased diversification. In addition, this prime had a larger influence on those with the

lowest levels of financial literacy — presumably those who were least informed about how to appropriately allocate their investments. Thus, shifting beliefs about epistemic and aleatory market uncertainty may prove to be a useful intervention for improving financial decision making among those who need it the most.

Chapter 3: The Influence of Forgotten Features on Product Choice

ABSTRACT

Consumers must often make product choices based on information contained in memory where some features are forgotten and some features are remembered. Across five studies, we show that choices are influenced by inferences consumers make about these forgotten features. In Study 1, we show that when consumers remember positive product features they tend to draw overly positive inferences about forgotten features. We then show how making overly positive inferences about forgotten features can bias consumers to choose a product from memory over an equivalent fully described product in the domains of gift cards, movie collections, and TVs in Studies 2, 3 and 4. Conversely, we show that when consumers make overly negative inferences about forgotten features they tend to reject products from memory in Study 5.

Many consumption decisions rely on memory. Consider, for example, the following scenario: On an initial shopping trip, a consumer learns about the features of a TV based upon a complete set of information. Later, at another store, the consumer examines other TVs. However, this store does not carry some of the previously examined brands. The consumer has complete, externally available information on all TV features in the second store, but for brands examined earlier, some features may have been forgotten. In such situations, it is unclear how a consumer would account for forgotten features when choosing between brands.

Much prior work has attempted to understand how consumers make product choices from memory. However, none of this research has focused on how inferences about features that have been learned, and then forgotten impact choice. Past work has primarily suggested that consumers focus on information that they remember, and treat forgotten information as a source of uncertainty (e.g., Dick, Chakravarti, and Biehal 1990). When making choices between products from memory, this tendency has been shown to lead consumers to make comparisons between remembered attributes (Biehal and Chakravarti 1983), and to prefer to include a more completely remembered product in a choice set (Nedungadi 1990). More generally, past research has shown that consumers tend to prefer products that are more completely remembered because there is less uncertainty associated with these products (Biehal and Chakravarti 1982, 1983, 1986; Lynch and Srull 1982).

We propose that consumer choice is also influenced by inferences about information that has been *forgotten*. For example, when choosing between a TV held in memory, and a second TV that is fully described, a consumer may also realize that they have forgotten some of the features of the TV held in memory. The consumer then may make some inferences about the gist of these forgotten features. If these inferences are positive, we suggest a consumer might

prefer the TV held in memory, despite the greater uncertainty associated with this product, contrary to past research.

In what follows, we first suggest that consumers overestimate how diagnostic remembered information is of forgotten information. Following this logic, we propose that when consumers tend to remember more positive product features they will draw overly positive inferences about forgotten features whereas when consumers tend to remember more negative product features they will draw overly negative inference about forgotten features. We go on to suggest that these inferences can drive consumers to choose a product held in memory over an equivalent fully described product when remembered features are overly positive and to choose a fully described product over an equivalent product held in memory when remembered features are overly negative.

INFERENCES ABOUT FORGOTTEN PRODUCT FEATURES

Previous work has demonstrated that people are forgetful and often fail to recognize errors in memory. Although most information stored in long-term memory is said to be “available” (Lewis 1979), only a small fraction of available information is accessible at a given time (Tulving and Pearlstone 1966) and some information is lost altogether (Squire 1987). Memory becomes less accessible with the passage of time (Wixted and Ebbesen 1997), can diminish within seconds of learning information (Peterson and Peterson 1959), or can be inhibited by recalling related information (Postman and Underwood 1973). Forgetting also tends to occur because insufficient attention is devoted to a stimulus at the time of encoding (Reason

and Mycielska 1982), attention is divided (Craik et al. 1996), or because attended information is processed superficially (Craik and Lockhart 1972).

Despite this tendency to forget, little research has focused on how consumers draw inferences about the qualities of forgotten product features. However, substantial research has investigated how consumers draw inferences about missing attribute values. Here, it is important to draw a distinction between forgotten features and missing attributes. For example, a TV feature can be forgotten (e.g., a consumer learns, then forgets that a TV has a 3D feature) or an attribute can be missing (e.g., a TV description contains the 3D TV feature, but contains no information about the quality of this feature). Past research has focused almost entirely on how consumers draw inferences about missing attributes (e.g., Broniarczyk and Alba 1994; Dick, Chakravarti, and Biehal 1990; Meyer 1981; Sanbonmatsu, Kardes, and Herr 1992) rather than forgotten features.

In this paper we propose consumers might also draw inferences about forgotten features. The literature on missing attributes has found that consumers primarily draw three types of correlational inferences about missing attributes based on known attributes. First, in a process known as evaluative consistency, consumers examine known attributes to infer an overall impression of a product or brand and assume that missing attributes adhere to this impression (Beckwith and Lehmann 1975; Broniarczyk and Alba 1994; Cooper 1981; Dick, Chakravarti, and Biehal 1990; Nisbett and Wilson 1977). Second, in a similar but more complex process, known as probabilistic consistency, consumers examine a known attribute and infer a missing attribute level based on prior beliefs about the correlation between attributes (Downing, Sternberg, and Ross 1985; Ford and Smith 1987; Hoch 1984; Huber and McCann 1982; John, Scott, and Bettman 1986; Kardes and Sanbonmatsu 1993; Meyer 1981; Ross and Creyer 1992).

For example, a TV with a good warranty will also be inferred to be durable based on the belief that higher quality warranties are associated with greater durability. Third, consumers might also make a compensatory inference where higher attribute levels for one feature indicate lower levels on another feature (Bradlow, Hu, and Ho 2004; Chernev and Carpenter 2001; Simonson, Carmon, and O'curry 1994). This is based on the implicit assumption that markets are efficient and products must sacrifice high performance on one feature to achieve high performance on another feature. For example, a product that is affordable might be expected to be low in durability.

In the case of forgotten features, we expect consumers to use the principles of evaluative consistency where they draw correlated inferences about the qualities of forgotten features based on an overall impression of the product or brand. We predict that consumers will be less likely to make inferences that are probabilistically consistent or compensatory when assessing forgotten features. To make this assumption, we draw on past research that shows inferences based on probabilistic consistency or compensatory inference are made when a consumer can apply a belief about the relationship between two attributes (where both features are known and one of the attribute levels is unknown) (e.g., Bradlow, Hu, and Ho 2004; Chernev and Carpenter 2001; Ross and Creyer 1992). In contrast, when a feature has itself been forgotten, a consumer has less meta-information about the relationship between the forgotten feature and other known feature. Instead, we expect a consumer will tend to make a higher-level inference, as is the case with evaluative consistency. We predict, in line with evaluative consistency:

H1: Consumers will draw positively correlated inferences about the quality of forgotten features based on the quality of remembered features.

BIASES IN MEMORY

Consumers may be biased to remember more positive product information in some cases whereas in other cases consumers may remember product information that is more negative. Past research shows that consumers tend to better encode information that is consistent with processing goals at the time of exposure (Bettman 1979; Keller 1987), which could bias consumers to remember more positive or negative information based on these goals. For example, when a consumer is choosing a product they may focus disproportionately on positive features whereas when a consumer is rejecting a product they may focus on negative features (Shafir 1993; Shafir, Simonson, and Tversky 1993; Shen and Wyer 2008). Consumers in good vs. bad moods may recall more positive vs. negative product features (Matt, Vázquez, and Campbell 1992) or consumers with a promotion vs. prevention regulatory focus may encode more positive vs. negative features, respectively (Higgins 1998; Lockwood, Jordan, and Kunda 2002; Wang and Lee 2006; Werth and Foerster 2007). Consumers also tend to more readily remember brand names that convey meaning (Keller, Heckler, and Houston 1998), or familiar brands which are often rated as more favorable (Kent and Allen 1994). Moreover, initially remembered information can diminish the ability to recall subsequent product features (Burke and Srull 1988), exacerbating these effects. In the domain of eye witness testimony people tend to have stronger memories for more extreme negative events (Burke, Heuer, and Reisberg 1992; Christianson 1992; Heuer and Reisberg 1990). However, the details of these memories are often less accurate than for more mundane events (Christianson and Nilsson 1984; Clifford and Hollin 1981; Clifford and Scott 1978; Loftus and Burns 1982). Advertisers might also influence

consumer memory by showing more positive information first or last to induce primacy or recency effects (Kahneman et al. 1993; Li 2010; Peters and Bijmolt 1997). More generally, research finds that more atypical, positive and negative events tend to be more memorable than everyday occurrences (Brown and Kulik 1977; Dutta, Kanungo, and Freibergs 1972; Hastie and Kumar 1979; Ochsner 2000).

But past research suggests consumers are often unaware of these recall biases. For example, when people were asked to recall a time that their preferred team won a football game, they tended to recall the most positive instance and make overly positive forecasts of their affective reaction to future events because they were unaware that they were recalling a biased sample of information (Morewedge, Gilbert, and Wilson 2005). In the same paper, when participants were asked to recall a time they missed the subway they tended to recall the most negative instance and make overly negative forecasts of their future affective reactions to similar events. Similarly, people tend to overgeneralized conclusions from atypical samples to wider populations (Hamill, Wilson, and Nisbett 1980). In this paper, we build on these prior findings to predict that:

H2: When consumers tend to selectively remember overly [positive/negative] product features consumers will draw overly [positive/negative] inferences about forgotten product features.

CHOICES FROM MEMORY

We predict that inferences about forgotten features will impact choices between a product that is held in memory and a product that is fully described. Substantial past research has found that consumers often prefer a fully described product over an equivalent product from memory (Biehal and Chakravarti 1982, 1983, 1986; Lynch and Srull 1982) or a product that is only partially described (Kivetz and Simonson 2000). This research has attributed this advantage to consumers choosing the option about which they have greater confidence and less uncertainty (Dick, Chakravarti, and Biehal 1990), consistent with the uncertainty effect (Gneezy, List, and Wu 2006). However, other research has found preferences for products from memory is very context dependent (Alba, Marmorstein, and Chattopadhyay 1992), and consumers tend to choose products with incomplete information when the information that is available is positive (Sanbonmatsu et al. 2003). Consumers explicitly prompted to make inference also tend to be more likely to choose a product with incomplete information over a product that is fully described (Gunasti and Ross 2009). In the present paper, we propose that the choice between a product from memory and a product from description also depends on whether consumers draw overly positive or negative inferences about forgotten features on the product from memory. We predict:

H3: When a consumer remembers positive features, and draws overly positive inferences about missing features, choice share will increase for a product from memory over an equivalent product from description

H4: When a consumer remembers negative features, and draws overly negative inferences about missing features, choice share will decrease for a product from memory over an equivalent product from description

Our research thus aims to address an important question that has been overlooked by prior literature: what inferences do consumers make about product features that have been forgotten and how do these inferences impact product choices from memory?

Across five studies, we examine how consumers make choices between products from memory and products from description. In Study 1, we examine a multi-retailer gift card and demonstrate that consumers draw overly-extreme inferences about forgotten retailers based on the attractiveness of remembered retailers (H1 and H2). We go on to find that when consumers remember overly positive feature they tend to choose a real product from memory over an equivalent real product from description in the domains of gift cards in Study 2, movie collections in Study 3, and TVs in Study 4 (H3). In Study 5 we examine product reviews and show that when consumers remember overly negative reviews they will tend to reject a TV from memory and instead choose a TV where the reviews are externally available (H4).

STUDY 1

In Study 1, we examined multi-retailer gift cards to test the hypothesis that consumers will draw overly-extreme inferences about forgotten features from remembered features (H1 and H2). In this study, we conceptualized the different retailers on the gift card as different “features.” We expected that consumers would tend to remember retailers that they consider

higher in quality and more familiar based on past findings (Kent and Allen 1994). We also expected that consumers would remember higher quality retailers based on the processing goal of using the card at high quality retailers (Bettman 1979; Keller 1987). We confirmed that consumers remembered retailers rated as higher quality in a pretest. Thus, we predicted participants would make overly positive forecasts of the quality of the retailers that have been forgotten (H1 and H2).

Method

We recruited 506 participants (55.7% female, mean age = 35.4, $SD = 12.1$ years) on Amazon Mechanical Turk to participate in an experiment in exchange for a \$.50 payment.

Participants recruited to the study were told they would be evaluating a gift card and given the following instructions:

“You will first learn about a gift card. You will be shown all of the retailers at which the gift card can be used one at a time. You will see each retailer for 3 seconds, then the page will automatically advance to the next retailer.

Please pay attention as you will have to answer questions about this gift card.”

Participants then completed a learning phase where they were shown each of the retailers, one at a time. All participants viewed 30 retailers, randomly selected from a list of 60 retailers (see Appendix B) in a random order for 3 seconds each. As a filler task, participants then provided demographic information.

Remembered Retailers. Participants were next asked to recall as many of the retailers as they could in a text box.

Remembered Retailer Quality. The recalled retailers were next reported back to the participants. Participants provided the rated *remembered retailer quality*, following the prompt: “Consider the average quality of the retailers you were **able to recall**, above. Please rate the average quality of these retailers.” Participants rated the average quality of the remembered retailers using a 0-100 slider where 0 was lowest quality and 100 was highest quality. On the next page, participants then answered two demographic questions, “Is English your first language? Y/N” and “Are you fluent in English? Y/N”. The purpose of placing these questions here was to separate the *rated quality of remembered retailers* from the *estimated forgotten retailer quality* so that we could minimize carryover effects from responding on the first question to responding on the second question.

Estimated Forgotten Retailer Quality. On the next page, participants estimated the quality of forgotten retailers by rating “the average quality of the retailers on the gift card you were **unable to recall** (retailers you did not list above, but were still on the gift card)” on the same 0-100 scale.

Estimated Number of Forgotten Retailers. On the next page, participants estimated the total number of retailers on the gift card in a numeric text box. We calculated estimated forgotten retailers by subtracting the number of remembered retailers from the estimated number of retailers.

Identifying Forgotten Retailers. Participants were then provided with the list of retailers they recalled and a master list of all of the retailers. Participants were asked to check off each of the remembered retailers from the master list to attain a list of forgotten retailers.

Actual Forgotten Retailer Quality. Participants were then asked to rate the quality of each of the forgotten retailers individually, using the same 0-100 scale. These ratings were averaged to compute the actual forgotten retailer quality.

Participants were then debriefed.

Results and Discussion

Selective Forgetting. We first confirmed that participants would selectively forget less appealing features. Actual forgotten retailer quality ratings averaged 57.8, ($SD = 16.1$), less than remembered average retailer quality ratings of 67.6 ($SD = 19.9$), $t(505) = 11.92, p < .001$, but higher than the scale midpoint of 50, $t(505) = 10.81, p < .001$.

Estimating Forgotten Retailer Quality. We next examined the prediction that consumers would overestimate the quality of forgotten features based on more positive remembered retailers (H1 and H2). In line with this hypothesis, we found that participants estimated that the quality of forgotten retailers was 61.6 on average ($SD = 18.2$) whereas the actual rated quality was 57.8 on average ($SD = 16.2$), $t(505) = 5.32, p < .001$.

To further test our prediction, we examined if participants drew overly extreme inferences from the quality of remembered retailers to estimate the quality of forgotten retailers. We predicted that *remembered retailer quality* ratings would be more predictive of *estimated forgotten retailer quality* than it is of *actual forgotten retailer quality*. To test this we ran two regressions with *remembered retailer quality* as the independent variable in both regressions and *estimated forgotten retailer quality* as the dependent variable in the first regression and *actual forgotten retailer quality* as the dependent variable in the second regression. *Remembered retailer quality* was more predictive of *estimated forgotten retailer quality*, $b = .46, CI = [.40, .54], p < .001$ than it was of *actual forgotten retailer quality*, $b = .39, CI = [.33, .45], p < .001$.

As predicted, the difference between these coefficients was significant $\chi^2(1,505) = 4.04, p < .05$, indicating that participants drew overly extreme inferences from remembered retailers when estimating forgotten retailer quality.

Number of Forgotten Retailers. We also examined how accurately participants assessed how many retailers they were forgetting. Participants underestimated that they forgot 18.8 retailers on average ($SD = 4.9$), when in fact they forgot 20.8 retailers on average ($SD = 13.9$), $t(505) = 3.56, p < .001$.

In sum, participants drew strong inferences about the quality of retailers they forgot from retailers they remembered. As predicted, participants remembered the retailers rated higher in quality and showed a tendency to overestimate how closely the quality of forgotten retailers resembled the remembered retailers. Said another way, participants were unaware of the degree to which they tended to forget less appealing retailers and remembered more appealing retailers.

STUDY 2

In Study 1, we found some evidence for our prediction that consumers would draw overly positive inferences about forgotten features when remembered features were positive (H1 and H2). However, this was in a stylized experiment and it's not clear whether inferences about forgotten features influence real choices. In Study 2, we directly examined how inferences about forgotten retailers would influence choices between a gift card that is remembered ("gift card from memory") and an equivalent gift card that is fully described ("gift card from description") in an incentive-compatible experiment. We predicted that participants would remember the higher quality retailers, overestimate the quality of the forgotten retailers on the card from

memory (H1 and H2) and choose the card from memory over the card from description (H3). We consider this a conservative test for three reasons. First, consumers show preferences for products where all attributes are known compared to products where some attribute levels are unknown (Biehal and Chakravarti 1982, 1983, 1986; Lynch and Srull 1982). Second, the uncertainty effect suggests participants tend to prefer a gift card that has a fixed value, over a gift card where the value is uncertain (Gneezy, List, and Wu 2006). Third, in Study 1 we found that participants underestimated the number of retailers they forgot, so the card from memory may be perceived to have fewer retailers and be less fungible.

We tested our prediction that participants would prefer a gift card from memory in a within-subject experiment where participants learned about a gift card, as in Study 1. Next, Participants chose between a gift card from memory and a comparable gift card from description where all of the retailers are fully described at the time of choice. In the present study, we used an incentive-compatible design, in which participants were told that some participants would be randomly selected to receive their actual gift card choice.

Method

We recruited 289 participants (43.0% female; mean age = 32.9; $SD = 10.1$ years) on Amazon Mechanical Turk were recruited to complete an experiment in exchange for a \$.60 payment.

We presented participants with two gift cards. Gift cards could be used at 30 retailers randomly generated from a list of 90 retailers (Appendix C) and were named either the “Eagle Gift Card” or the “Sapphire Gift Card” at random. One card was randomly selected to be the “gift card from memory” and one was selected to be the “gift card from description”. Participants first learned about the 30 retailers on the gift card from memory following the same procedure

from Study 1. That is, participants viewed each of the 30 retailers for 3 seconds each in a randomized order.

Average Retailer Quality. Participants were then provided with a complete list of the 30 retailers on the fully described gift card. On this same page, participants were asked to “Consider the average quality of the retailers offered on both cards. Please rate which card has a higher average retailer quality.” Participants rated which card had higher average retailer quality on a 6-point scale where 1 or 6 was randomly assigned to be labeled as “definitely the Eagle Gift Card” and the opposite scale end-point was labeled “definitely the Sapphire Gift Card.” We then recoded the average retailer quality to be 1 when the Eagle Gift Card was maximally preferred and 6 when the Sapphire Gift Card was maximally preferred.

Card Choice. On the same page participants then selected which card they would prefer following the prompt:

“A few participants will be selected at random to receive a \$25 gift card.

Think back to the [Eagle/Sapphire] Gift Card (the first card you viewed) and compare it to the [Eagle/Sapphire] Gift Card (listed above). If you are selected would you prefer to receive the [Eagle/Sapphire] or the [Eagle/Sapphire] Gift Card? This choice will be binding if you are drawn so choose carefully.”

Participants then chose between the two gift cards, which were presented below in a random order. Participants also responded to the open prompt, “What factors determined your gift card choice?”

Estimated Number of Retailers. On the following page participants estimated the number of retailers on the gift card from memory in a numeric text box.

Participants then provided demographic information and completed a short attention check and were debriefed.

Results and Discussion

Average Retailer Quality. We first tested our prediction that participants would tend to estimate the average quality of features on the card from memory as higher than on the card from description. To do so, we coded the condition as 0 when the Eagle Gift Card was from memory and the Sapphire Gift Card was from description and 1 when the Sapphire Gift Card was from memory and the Eagle Gift Card was from description. As described in the methods, average retailer quality of 6 indicates the Sapphire Gift Card is perceived to have higher quality and 1 the Eagle Gift Card is perceived to have higher quality. We then regressed average retailer quality onto experimental condition. As predicted, we found that when the gift card was forgotten, it was rated as higher in average retailer quality, $b = .58$, 95% CI = [.23, .93], $p < .001$.

Gift Card Choice. We next examined our prediction that participants would choose the gift card from memory over the gift card from description. Confirming our prediction, 57.1% of participants preferred the card from memory, (binomial test, 165 out of 289, test value .5, $p < .05$).

Mediation of Gift Card Choice. We next examined whether participants chose the gift card from memory because they overestimated the quality of forgotten retailers from remembered retailers (H3). To do this we conducted a bootstrap mediation analysis for dichotomous outcomes, in which average retailer quality was the mediator between condition and gift card choice (see Figure 5). We used average retailer quality as an inferential measure of the quality of forgotten retailers since this measure captures expectation about the quality of forgotten and remembered retailers. First, we found that experimental condition was a reliable

predictor of gift card choice, $b = .55$, 95% CI = [.09, 1.02], $p < .05$; i.e., the gift card that had forgotten features was a reliable predictor of final gift card choice. Second, we found that condition was a reliable predictor of average retailer quality, $b = .58$, $p < .001$, and average retailer quality predicted gift card choice, $b = 1.13$, 95% CI = [.89, 1.38], $p < .001$. Third, when both condition and average retailer quality were included in the model, we found that average retailer quality predicted gift card choice, $b = 1.15$, 95% CI = [.90, 1.40], $p < .001$, whereas experimental condition was no longer a reliable predictor of choice, $b = .19$, 95% CI = [-.42, .79], $p = .544$. Overall, we found that average product quality explained 78.8% of the experimental treatment effect on gift card choice, $b_{\text{indirect}} = .13$, 95% bootstrapped 95% CI = [.05, .21] (bootstrapping based on 500 resamples).

Number of Forgotten Retailers. Replicating results from Study 1, participants estimated that there were 26.3 retailers on the gift card from memory on average ($SD = 9.9$), whereas the gift card actually contained 30 retailers, a significant underestimation, $t(288) = 6.30$, $p < .001$.

An alternative explanation for why average retailer quality was rated higher on the card from memory is that participants underestimated how many retailers were forgotten. Thus, in calculating the average retailer quality, participants may have over-weighted the remembered retailers and under-weighted the forgotten retailers, resulting in an overestimate on the card from memory. One way to test for this possibility is to examine if average retailer quality is higher when the estimated number of forgotten retailers is lower. To do so, we recoded average retailer quality to be 6 when the card from memory was rated as highest and 1 when the card from description was rated as highest. We then regressed average retailer quality on number of forgotten retailers. Thus, if this alternative explanation were true, one would expect a higher number of forgotten retailers to predict a lower rating. However, we found no support for this

alternative explanation: a higher number of forgotten retailers was related to *higher* average retailer quality, $b = .019$. 95% CI = [.002, .038], $p < .05$.

In sum, we found data consistent with the hypothesis that participants remembered the higher quality retailers and overestimated the quality of the forgotten retailers on the card from memory (H1 and H2). This led to participants choosing the card from memory over an equivalent card from description (H3). We found participants chose the card from memory even though they believed that there were fewer retailers on this gift card. This result is also notable since prior work has shown that consumers generally prefer products from description over products from memory (e.g., Biehal and Chakravarti 1982, 1983, 1986; Lynch and Srull 1982).

STUDY 3

In study 3 we address some shortcomings of Study 2. First examine our prediction in the domain of movie package. The context of a multi-retailer gift cards used in Study 2 may be especially prone to memory errors since, first, consumers may not be familiar with this type of product. Second, since the gift card is only worth \$25 a consumer may plan to only use the card at fewer than 5 stores and therefore see no reason to remember more than 5 stores. In contrast, consumers are familiar with purchasing collections of movies with many popular services offering movie collections such as Amazon Prime, Netflix, Hulu, and the Criterion Collection. Watching one movie also does not eliminate the options to watch another movie as is the case with the gift card. We also replaced the likert measure of quality with a dichotomous choice. We addressed possible order effects in Study 1 by having participants make the quality judgment

after the choice on a separate page, whereas in the gift card study it was before the choice on the same page.

In Study 3 we replicated the results of Study 2 in the context of movie collections where participants made choices between a movie collection from memory and an equivalent movie collection from description. We conceptualized different movies in the collection as “features” of the movie collection. As in Studies 1 and 2, we expected that participants would remember movies that they consider to be higher in quality due to familiarity (Kent and Allen 1994) and processing goals (Bettman 1979; Keller 1987) and forget movies lower in quality. We predicted participants would overestimate the quality of the forgotten movies on the movie collection from memory (H1 and H2) and choose the collection from memory over the collection from description (H3).

Method

We recruited 409 participants (48.2% female; mean age = 37.1; $SD = 12.1$ years) on Amazon Mechanical Turk to complete an experiment in exchange for a \$.60 payment.

The methods closely followed those in Study 2, except with collections of movies rather than gift cards. We presented participants with two movie collections. Each collection was composed of 30 movies randomly generated from a list of the 60 greatest movies of all time as rated by IMDB (Appendix D) and were named either the “Brighton Collection” or the “Oxford Collection” at random. One collection was randomly selected to be the “movie collection from memory” and one was selected to be the “movie collection from description”.

After a brief introduction, participants were first told:

“In this survey you will learn about and evaluate two collections of movies: the [Brighton/Oxford] Collection and the [Brighton/Oxford] Collection.

You will first learn about the [Brighton/Oxford] Collection. You will be shown all of the movies in the [Brighton/Oxford] Collection one at a time. You will see each movie for 3 seconds, then the page will automatically advance to the next movie.

Please pay attention as you will have to answer questions about this collection.”

Participants then learned about the 30 movies on the collection from memory following the same procedure from Study 3. That is, participants viewed the title of each of the 30 movies for 3 seconds each in a randomized order.

Collection Choice. Participants were then provided with a complete list of the 30 movies in the movie collection from description. Participants were told that some participants would be selected to receive their real choice. Participants then followed the prompt below to choose between the collections (the position of the radio buttons for the choices were randomized):

“We want you to consider whether you would rather receive all the movies in the [Oxford/Brighton] Collection or the [Oxford/Brighton] Collection.

You would receive all of the movies contained in the chosen collection and could watch these movies anytime, so consider how much you value having access to all of the movies.

Think back to the [Oxford/Brighton] Collection (the first collection you viewed) and compare it to the [Oxford/Brighton] Collection (listed above). If you are selected would you prefer to receive the full [Oxford/Brighton] Collection or the full [Oxford/Brighton] Collection?”

Quantity Judgment. On the next page participants were asked, “Which of the collections has more movies?” Participants then responded in a dichotomous choice. The list of movies on the movie collection from description was shown at the top of the page.

Quality Judgment. On the same page as the quantity judgment participants were asked to indicate “Which of the collections has higher quality movies, on average?” Participants then responded in a dichotomous choice. Participants then provided demographic information and completed a short attention check and were debriefed.

Results and Discussion

Movie Quality. We first tested our prediction that participants would tend to believe the quality of movies was higher on the collection from memory than on the collection from description (H1 and H2). As predicted, 60.9% of participants believed the movie collection from memory had higher quality movies, (binomial test, 227 out of 408, test value .5, $p < .001$).

Movie Collection Choice. We examined which movie collection participants would choose. As predicted, and replicating Study 2, 55.5% of participants preferred the collection from memory, (binomial test, 165 out of 289, test value .5, $p < .05$). As a further test, we examined whether collection choice was more related to movie quality or quantity. To do so, we conducted a logistic regression with collection choice as the dependent variable and judgments of movie quality and movie quantity as independent variables. As predicted, movie quality predicted choice, $b = 3.81$, 95% CI = [3.19, 4.42], $p < .001$, whereas movie quantity was not related to choice, $b = .01$, 95% CI = [-.55, .58], $p = .971$.

Mediation of Collection Choice. We next examined whether participants chose the movie collection from memory because they overestimated the quality of forgotten movies from remembered movies (H3). To test this we conducted a bootstrap mediation analysis for

dichotomous outcomes, in which movie quality was the mediator between within-subject condition and collection choice. First, we found that experimental condition was a reliable predictor of collection choice, $b = .44$, $CI = [.05, .83]$, $p < .05$; i.e., the collection that was from memory was a reliable predictor of final collection choice. Second, we found that condition was a reliable predictor of movie quality, $b = .88$, $p < .001$, and collection quality predicted collection choice, $b = 3.68$, $CI = [3.12, 4.25]$, $p < .001$. Third, when both condition and movie quality were included in the model, we found that movie quality predicted choice, $b = 3.82$, $CI = [3.21, 4.42]$, $p < .001$, whereas experimental condition was no longer a reliable predictor of choice, $b = -.45$, $CI = [-1.05, -.16]$, $p = .149$. In a bootstrap model we found the indirect effect was significant, $b_{\text{indirect}} = .17$, 95% bootstrapped $CI = [.10, .25]$ (bootstrapping based on 500 resamples).

Movie Quantity. Replicating results from Studies 1, 2 and 3, 55.6% of participants believed the movie collection from memory had fewer movies, (binomial test, 227 out of 408, test value .5, $p < .05$).

A possible alternative explanation for why average movie quality was rated higher on the collection from memory is that participants underestimated how many movies were forgotten. Thus, in calculating the average movie quality, participants may have over-weighted the remembered movies and under-weighted the forgotten movies, resulting in an overestimate on the card from memory. One way to test for this account is to examine if the collection from memory is less likely to be judged as having higher quality movies when participants think it has more movies than the collection from description. Since participants viewed the collection from description while making these judgments, the collection from memory should only be judged to have more movies if participants overestimate how much they are forgetting. We examined this possibility in a logistic regression with movie quality as the dependent variable and movie

quantity as the independent variable. We found no support for the alternative explanation: participants that judged the collection from memory to have more movies also tended to judge that the collection from memory had higher quality movies, $b = .38$, 95% CI = [-.03, .78], $p = .067$.

In sum, we found evidence to suggest that participants remembered the higher quality movies, overestimated the quality of movies that had been forgotten, and overestimated the quality of movies in the collection from memory (H1 and H2). This led to participants choosing the collection from memory over the collection from description (H3). These findings replicate Study 2 while extending them to the domain of movie collection, and using different measures of subjective value. We also found that participants believed there were fewer movies on the collection from memory compared to the collection from description. This finding suggests that positive inferences based on remembered retailers were only extended to judgments of quality, and not to more general judgments about the number of movies in the package.

STUDY 4

In Study 4, we addressed limitations of the prior studies in several ways. First, we extended our findings to TV features to show the generalizability of our findings. Both multi-retailer gift cards and movie collections are products where each of the features (i.e., retailers and gift cards) are easily split. This may have led to greater forgetting of less appealing features in these prior studies.

Second, we extended our measures beyond feature quality to include an overall evaluation of the total value of the TV features. Measuring the total value of TV features

addresses an important alternative explanation for our prior findings. In our prior studies, we examined average feature quality as our primary indicator of the inferences participants made about forgotten features. However, average feature quality may be inflated when participants underestimate the number of features they forget. To resolve this issue, we asked participants to assess the total value of features in this study. We also included a measure of confidence since prior work shows that consumers often prefer products from description over products from memory due to low confidence (e.g., Biehal and Chakravarti 1982, 1983, 1986; Lynch and Srull 1982).

In this study we extended the results of Studies 2 and 3 in the context of a real TV choice where participants made choices between a TV from memory and an equivalent TV from description. As in Studies 1, 2, and 3, we expected that participants would remember TV features that they considered to be most valuable because a consumer might have the processing goal of remembering these features (Bettman 1979; Keller 1987). We predicted participants would overestimate the value of forgotten features on the TV from memory (H1 and H2) and choose the TV from memory over the TV from description (H3).

Method

We recruited 193 participants (45.1% female; mean age = 35.2; $SD = 11.1$ years) on Amazon Mechanical Turk to complete an experiment in exchange for a \$.60 payment plus the possibility of winning a TV.

The methods closely followed those in Studies 2 and 3, except with TVs. We presented participants with two TVs. Each TV was described as a “60-inch HD TV” and was composed of 25 features generated from a list of 50 real features on a Samsung TV (Appendix D). TVs were labeled a “Samsung TV” or a “Sony TV” and were randomly selected to be either the “TV from

memory” or the “TV from description”. Participants first learned about the 25 features on the TV from Memory collection following the same procedure from Study 3. That is, participants viewed each of the 25 features for 3 seconds each in a randomized order.

Average Feature Quality. Participants were then provided with a complete list of the 25 features on the TV from description. Participants were asked to “Consider the average quality of the features offered on both TVs. Please rate which TV has higher quality features, on average.” Participants rated which TV had higher average feature quality on a 6-point scale where 1 or 6 was randomly assigned to be labeled as “definitely the Samsung TV” and the opposite scale end-point was labeled “definitely the Sony TV.” We then recoded the feature quality to be 1 when the Samsung TV was maximally rated and 6 when the Sony TV was maximally rated.

TV Choice. On the same page participants then chose between the TVs (the position of the radio buttons for the choices were randomized):

“A participant will be selected at random to receive their actual TV choice. Winners will be notified and the TV will be shipped direct via amazon.

Think back to the [Samsung/Sony] TV (the first TV you viewed) and compare it to the [Samsung/Sony] TV (listed above). If you are selected would you prefer to receive the [Samsung/Sony] TV or the [Samsung/Sony] TV? This choice will be binding if you are drawn so choose carefully.”

Total Feature Value. On the next page participants were asked to “Consider the total value of the features offered on both TVs. Please rate which TV has features that are more valuable to you, in total.” Participants rated which TV had higher average feature quality on a 6-point scale where 1 or 6 was randomly assigned to be labeled as “definitely the Samsung TV”

and the opposite scale end-point was labeled “definitely the Sony TV.” 1 when the Samsung TV was maximally rated and 6 when the Sony TV was maximally rated.

Estimated Number of Features. On the next page participants estimated the number of features on the TV from memory in a numeric text box.

Confidence. Finally, Participants assessed their confidence in their judgments by responding to the prompt: “How confident are you in your judgments about the [TV from memory]?” on a 1-7 likert scale where 1 is “not at all confident” and 7 is “extremely confident”.

Participants then provided demographic information and were debriefed.

Results and Discussion

TV Choice. We first examined which TV participants chose. As predicted, and replicating Studies 2 and 3, 61.2% of participants chose the TV from memory, (binomial test, 118 out of 193, test value .5, $p = .001$).

Average Feature Quality. We next tested our prediction that participant’s tendency to overestimate the quality of forgotten features would drive participants to perceived the features on the TV from memory to be higher in quality than the features on the TV from description (H3). To do so, we coded the condition as 0 when the Samsung TV was from memory and the Sony TV was from description and 1 when the Sony TV was from memory and the Samsung TV was from description. As described in the methods, average feature quality was recoded so that 0 indicates the Sony TV is perceived to have higher quality and 1 the Samsung TV is perceived to have higher quality. We then regressed average feature quality onto experimental condition. As predicted, we found that when the TV was from memory, it was rated as higher in average feature quality, $b = 1.08$, 95% CI = [.67, 1.49], $p < .001$.

Total Feature Value. We next tested our prediction that the tendency to overestimate the quality of forgotten features would drive participants to perceive the total value of features on the TV from memory to be higher than the TV from description (H3). To do so, we coded the condition as 0 when the Samsung TV was from memory and the Sony TV was from description and 1 when the Sony TV was from memory and the Samsung TV was from description. As described in the methods, average feature quality was recoded so that 6 indicates the Sony TV is perceived to have higher value and 1 the Samsung TV is perceived to have higher value. We then regressed total feature value onto experimental condition. As predicted, we found that when the TV was from memory, it was rated as higher in average feature quality, $b = 1.36$, 95% CI = [.93, 1.80], $p < .001$.

Estimated Number of Features. We next examined whether participants would underestimate how many features they forgot, as in Studies 1, 2, and 3. We found that participants underestimated that there were 20.7 features on the TV from memory on average ($SD = 8.0$), whereas the TV actually contained 25 features, a significant underestimation, $t(192) = 7.4$, $p < .001$.

Confidence. We then examined the relationship between confidence, choice, and quality judgment in a correlation analysis. For this analysis we recoded average feature quality and total feature value to be 6 when the TV from memory was maximally rated and to be 1 when the TV from description was maximally rated. As shown in Table 1, higher confidence in memory is correlated with choosing the TV from memory, rating the TV from memory as having higher quality features, and rating the TV from memory as having more valuable features.

Mediation of Average Feature Quality on TV Choice. To test the mediation model proposed in H3, we conducted a bootstrap mediation analysis for dichotomous outcomes, in

which average feature quality and total feature value were the mediators between experimental condition (whether the TV was from memory or description) and TV choice while controlling for confidence as a covariate (Figure 6). Confidence was recoded so that 1 is “not at all confident” and 7 was “extremely confident” when the Sony TV was from memory and the Samsung TV was from description; and 1 was “extremely confident” and 7 is “not at all confident” when the Samsung TV was from memory and the Sony TV was from description. First, as noted above, we found that experimental condition was a reliable predictor of collection choice, $b = .64$, 95% CI = [.03, 1.25], $p < .05$; i.e., the TV from memory was more likely to be chosen after controlling for confidence. Second, we found that condition was a reliable predictor of average feature quality, $b = .90$, $p < .001$ and total feature value, $b = 1.09$, $p < .001$ after controlling for confidence, i.e., the TV from memory was rated as having higher quality and more valuable features after controlling for confidence. Third, when condition, average feature quality, total feature value and confidence are included in the model, we found that average feature quality predicted choice, $b = 1.43$, 95% CI = [.81, 2.07], $p < .001$ and total feature value predicted choice, $b = 1.25$, 95% CI = [.72, 1.79], $p < .001$ whereas experimental condition was no longer a reliable predictor of choice, $b = -.74$, CI = [-1.90, .43], $p = .214$ and confidence was no longer a reliable predictor of choice, $b = .18$, CI = [-.38, .75], $p = .520$. In a bootstrap model we found the indirect effect was significant, $b_{\text{indirect}} = .30$, 95% bootstrapped 95% CI = [.19, .42], $p < .001$ (bootstrapping based on 500 resamples).

In sum, we found that participants overestimate the total value of the features on the TV from memory (H1 and H2) and choose the TV from memory over the TV from description (H3). These findings replicate Studies 2 and 3 while extending them to the domain of TV features. We also utilize a new measure of subjective value to infer the value of forgotten features. Finally, we

control for confidence. We found that participants who were more confident in their memories tended to choose the TV from memory more often, consistent with past research (e.g., Biehal and Chakravarti 1982, 1983, 1986; Lynch and Srull 1982). However, controlling for confidence did not significantly weaken our findings, suggesting judgments of confidence and forgotten feature value both contribute to product decisions from memory.

STUDY 5

In Study 5 we addressed an important issue in Studies 1-4. In Studies 1-4 we showed how consumers prefer products from memory because they made overly positive inferences about forgotten product information. In this study we showed that consumers tend to reject products from memory when they make overly negative inferences about forgotten product information. This helps to show that our prior findings cannot be explained by a general positivity bias for products from memory. We also added a new DV, “worst reviews” to show inferences about forgotten information impact judgments outside of quality (Studies 1-4) and value (Study 4).

In Study 5 we extended the results of Study 4 in a context where participants are presented with negative TV reviews from Amazon.com and make choices between a TV from memory and an equivalent TV from description. We expected that participants would remember TV reviews that they considered to be most negative because participants have the tendency to remember atypically extreme events (e.g., Brown and Kulik 1977; Dutta, Kanungo, and Freibergs 1972; Morewedge, Gilbert, and Wilson 2005; Ochsner 2000). We predicted participants would overestimate the negativity of forgotten reviews on the TV from memory (H1 and H2) and choose the TV from description over the TV from memory (H4).

Method

We recruited 196 participants (49.0% female; mean age = 38.8; $SD = 13.4$ years) on Amazon Mechanical Turk to complete an experiment in exchange for a \$.60 payment plus the possibility of winning a TV. In this study 10 (5.1%) participants failed a manipulation check and were excluded from analysis.

The methods closely followed those in Study 4, except participants were presented with negative TV reviews, rather than features. We presented participants with two TVs. Each TV was described as a “60-inch HD TV”. We then collected snippets from 50 real negative reviews on Amazon.com (Appendix E). To find these reviews we examined real reviews on Amazon.com for both Sony and Samsung 60-inch HD TV by searching for “Sony and Samsung TV” and examining the reviews of the first TVs to appear in the search results. We examined only reviews with 3 stars or less and selected 1-2 sentence snippets from these reviews. We selected snippets of reviews that we deemed to be (1) sufficiently negative and (2) expressing a different negative aspect than previously collected reviews. We examined approximately 100 negative reviews from approximately 10 different TVs to collect 50 negative reviews snippets. We then modified these review snippets to remove any brand identifying information. We then randomly assigned 25 of these review snippets to each TV. TVs were labeled the “Samsung TV” or the “Sony TV” and were randomly selected to be either the “TV from memory” or the “TV from description”. Participants first viewed 25 negative review snippets on the TV from Memory following the same procedure from Study 3. That is, participants viewed each of the 25 negative reviews for 3 seconds each in a randomized order.

Worst Reviews. Participants were then provided with a complete list of the 25 review snippets on the TV from description. Participants then answered the question: “Which of the TVs has the worst reviews, on average?” in a dichotomous choice question.

TV Choice. On the same page participants then chose between the TVs following the prompt:

“A participant will be selected at random to receive their actual TV choice. Winners will be notified and the TV will be shipped direct via amazon.

Think back to the [Samsung/Sony] TV (the first TV you viewed) and compare it to the [Samsung/Sony] TV (listed above). If you are selected would you prefer to receive the [Samsung/Sony] TV or the [Samsung/Sony] TV? This choice will be binding if you are drawn, so choose carefully.”

Demographics and Attention Check. Participants then provided demographic information. Participants also answered an attention check where they had to indicate which of the TVs they chose to receive in a multiple choice question. Participants were then debriefed.

Results and Discussion

TV Choice. We first examined which TV participants would choose. As predicted, and reversing Studies 2, 3 and 4, only 42.2% of participants chose the TV from memory, (binomial test, 79 out of 187, test value .5, $p < .05$).

Worst Reviews. We next tested our prediction that participants would rate the reviews as worse for the TV from memory. As predicted, and reversing Studies 2, 3 and 4, 58.3% of participants thought the TV from memory had worse reviews, (binomial test, 109 out of 187, test value .5, $p < .05$).

Mediation of TV Choice. To test the mediation model proposed in H4, we conducted a bootstrap mediation analysis for dichotomous outcomes, in which feature quality was the mediator between experimental condition (whether the TV was from memory or description) and TV choice. First, as noted above, we found that experimental condition was a reliable predictor of TV choice, $b = -.65$, 95% CI = [-1.25, -.07], $p < .001$; i.e., the TV from description was more likely to be chosen. Second, we found that condition was a reliable predictor of worse reviews, $b = .69$, $p < .05$, i.e., the TV from memory was rated as having worse reviews. Third, when both condition and feature quality are included in the model, we found that feature quality predicted choice, $b = -4.04$, 95% CI = [-4.97, -3.13], $p < .001$, whereas experimental condition was no longer a reliable predictor of choice, $b = -.31$, CI = [-1.23, .60], $p = .502$. In a bootstrap model we found the indirect effect was significant, $b_{\text{indirect}} = -.14$, 95% bootstrapped 95% CI = [-.26, -.02], $p < .05$ (bootstrapping based on 500 resamples).

In sum, we found that participants overestimate the severity of negative reviews on the TV from memory (H1 and H2) and choose the TV from description over the TV from memory (H4). These findings extended Studies 2, 3, and 4 by reversing the valence of inferences about forgotten information. We also showed the robustness of these effects by extending them into the domain of negative product reviews. These results have clear practical relevance as well. Given that product reviews are used to evaluate many different categories of products, and given that many products have hundreds of long reviews, it may be impossible to evaluate a specific product without to some extent cataloging the review information in memory.

GENERAL DISCUSSION

In this paper, we found that the evaluation of forgotten information is an important factor in consumer product judgments and choices from memory. In five studies, we found that consumers draw overly-extreme inferences about forgotten product features based on the characteristics of remembered features (H1 and H2). We went on to show that when consumers remembered overly positive features they tended to choose a real product from memory over an equivalent real product from description in the domains of gift cards in Study 2, movie collections in Study 3, and televisions in Study 4 (H3). In Study 5 we reversed these effects to show when consumers remembered overly negative product reviews they tended to reject a TV from memory and instead choose a TV where reviews are externally available (H4). In all studies involving real choices participants only made indirect evaluations of forgotten features, so as not to artificially induce participants to make inferences about forgotten features. This suggests that these inferences are automatic.

Another contribution of this work is to make a distinction between forgotten features and forgotten attributes. In the introduction, we noted that in this paper we examined forgotten features (e.g., if a person forgets that a smartwatch has GPS), whereas most past research on consumer inferences examines unknown attributes (e.g., if a person knows that a smartwatch has GPS, but doesn't know if the quality of this feature is good or bad). We found that participants made overly extreme inference about forgotten features, adhering to the principles of evaluative consistency. We predicted consumers would generally not make compensatory inferences since these types of inferences require some metaknowledge about the relationship between two features, and in our experimental paradigm the feature had been forgotten entirely. However, there could be a scenario where consumers make compensatory inference if the known feature is

correlated with *all* other features. For example, if a consumer remembers that a product is affordable, she may infer that all forgotten features are low in quality.

It should also be noted that these results may be dependent on time, as well as be affected by implicit memories about a brand. For example, Sanbonmatsu, Kardes, and Sansone (1991) found consumers made fewer inferences about missing attributes upon first learning about a product, but after a 1 week delay consumers made stronger inferences. In our studies participants made inferences with no delay suggesting that consumers might be more willing to make inferences about features that were forgotten, perhaps because these types of inferences are more natural. Lee (2002) finds that implicit memory based on prior knowledge of a brand also plays an important role in choices from memory. In our studies, we sought to limit any prior knowledge, and counterbalance all product information, so that prior brand knowledge and implicit memory would not be a factor. Future research should seek to investigate how implicit memories for brands might interact with inferences about forgotten features.

These studies also add to a body of research focused on understanding the effects of information presentation. Past studies have investigated the negative effects of information overload on consumers (Iyengar and Lepper 2000; Malhotra 1982; Scheibehenne, Greifeneder, and Todd 2010). This research finds more information can drive poor decision quality (Jacoby, Speller, and Kohn 1974) and less satisfied consumer decision making (Lee and Lee 2004). While we did not manipulate the amount of information in our studies, our findings nonetheless suggest that higher levels of information might lead to judgment and choice errors. As consumers are presented with many features they may be biased to selectively remember certain features and make incorrect inferences about forgotten features. Ultimately, when positive features are remembered, this may lead consumers to choose a lower quality, product from memory over a

higher quality product from description. In contrast, when negative features are remembered, consumers may reject a higher quality product from memory.

Although not the primary thrust of this paper, we also find that participants underestimated the number of features they forgot in Studies 1, 2, 3 and 4. This is consistent with a growing body of literature that shows that consumers tend to inaccurately assess the amount of missing information. First, in product judgments, consumers tend to focus on known information and ignore important omitted attributes when making judgments (Caputo and Dunning 2005; Sanbonmatsu, Kardes, and Herr 1992; Sanbonmatsu, Kardes, and Sansone 1991). In financial markets, investors tend to underreact to the absence of news after a merger (Giglio and Shue 2014), or the sudden silence of a company insider during an arbitration proceeding (Hong and Li 2015). Moviegoers also overestimate the quality of movies that a studio has chosen to withhold from critics, even though these movies tend to be of a much lower quality (Brown, Camerer, and Lovallo 2012). Similarly, people tend to ignore gaps in their knowledge and think they know the inner workings of various types of causal systems, from machines to public policies, in much greater detail than they actually do (Alter, Oppenheimer, and Zemla 2010; Fernbach et al. 2013; Rozenblit and Keil 2002). Consumers also tend to focus on product features for which attribute values are known and underweight features where attribute values are missing (Kivetz and Simonson 2000). More generally, people tend to focus on known evidence and underestimate the amount of unknown or missing evidence, leading to overconfidence (Walters et al. in press).

When taken together, these results can also be interpreted as a demonstration of overconfidence in the completeness and representativeness of remembered information when evaluating forgotten information. First, our finding that consumers make overly extreme inferences about forgotten information from remembered information implies that consumers are

overconfident in how representative remembered information is of forgotten information. Second, our finding that consumers *underestimate* how much has been forgotten implies that consumers are overconfident in the completeness of memory, and *overestimate* the proportion of information that has been remembered. This is consistent with past research that shows consumers are often overconfident in their prospective memory, underestimating the likelihood they will forget to claim a payment (Ericson 2011) or complete an experimental task (Guynn, McDaniel, and Einstein 1998). Failures of prospective memory also lead elderly patients to forget to take their medication (Park and Kidder 1996) and air traffic controllers to forget to take actions to direct aircraft (Vortac, Edwards, and Manning 1995). Similarly, a meta-analysis indicated that eye witness confidence in memory is only moderately correlated with accuracy (Bothwell, Deffenbacher, and Brigham 1987), again suggesting poor calibration in the assessment of memory. Future research might seek to investigate whether this bias towards overconfidence in remembered information found in our studies is related to overconfidence in prospective memory (Ericson 2011), and/or other domains of overconfidence (see Moore and Healy 2008). We hope our work will help consumers avoid these errors and improve decision making from memory.

Conclusion

Our investigation has shown that people draw inferences about information that is partially unknown to inform judgments and decisions. However, these inferences seem to be difficult to make and prone to errors. In Chapter 1 we showed that overconfidence was driven in part by peoples' tendency to underestimate the number of known unknowns. In Chapter 2 we showed how costly investment strategies such as day trading and active management were linked to the tendency to believe the that unknowns in the stock market are knowable (e.g., epistemic), whereas the academic community generally regards these unknowns as random (e.g., aleatory). In Chapter 3 we found that consumers made inferences about product features they had forgotten based on product features they remembered. However, these consumers failed to account for their tendency to remember biased information, leading to systematic errors in choices from memory.

This work also provides strategies for making more accurate inferences about unknowns. First, we found evidence in Chapter 1 that people with more complete knowledge structures make more accurate inferences about the amount of missing information. This is reflected in the famous quote by John Archibold Wheeler:

*We live on an island surrounded by a sea of ignorance. As our island of knowledge grows, so does the shore of our ignorance.*⁴

In line with this quotation, a larger knowledge structure will yield a wider shoreline. However, our research suggests that people can also improve their judgment by assessing the nature of these waters. Is the water shallow and knowable or deep and random? Is the island of

⁴ Scientific American (1992), Vol. 267

knowledge biased in any way by stronger memory for certain types of information? Our work shows that people take stock of this sea of ignorance when making decisions. I hope this work can also provide new tools for assessing these rough waters.

Tables

Table 1. Correlations between Chapter 3, Study 4 measures

	<i>M (SD)</i>	TV Choice	Feature Quality	TV Quality	Number Features	Confidence
TV Choice	.61 (.49)		.76***	.78***	.11	.22***
Feature Quality	4.0 (1.4)			.80***	.12	.19**
TV Quality	4.2 (1.5)				.15*	.27***
Number Features	20.7 (8.0)					.02
Confidence	3.4 (1.2)					

* $p < .05$, ** $p < .01$, *** $p < .001$.

Figures

Figure 1: Overconfidence, Chapter 1, Study 2

Confidence, percent correct, and overconfidence in the no treatment, consider the unknowns, and consider the alternative conditions. Confidence and percent correct are shown on the left vertical axis and overconfidence is shown on the right vertical axis. Standard errors displayed.

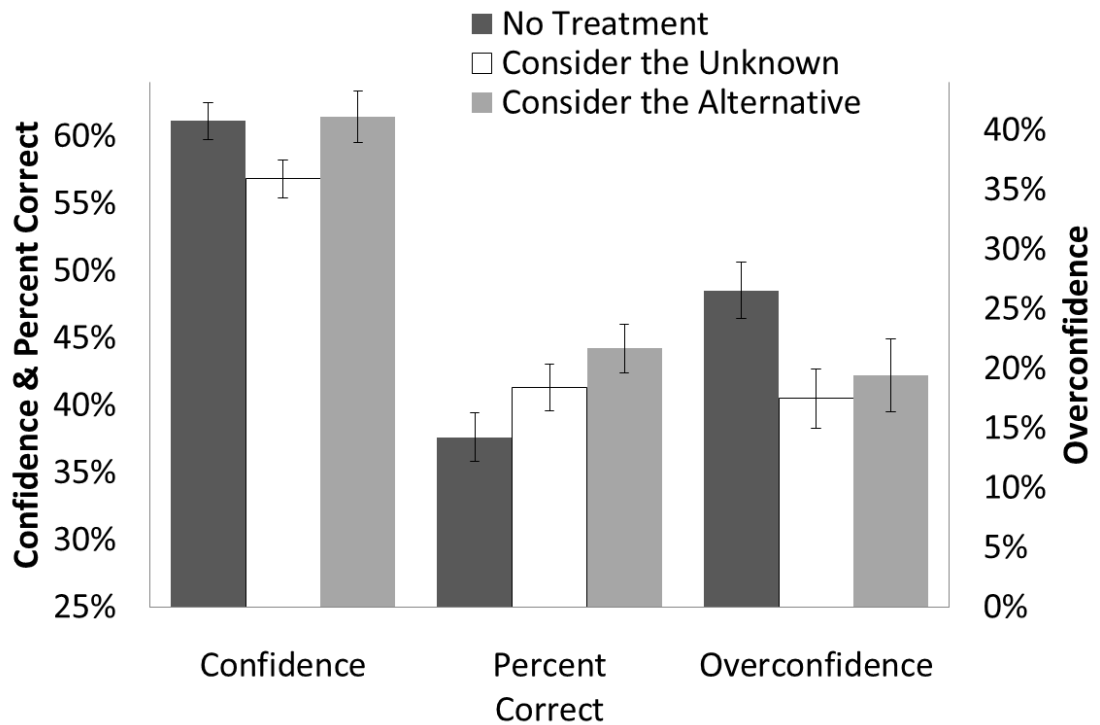


Figure 2: Effect of Consider the Unknowns, Chapter 1, Study 3

Confidence, percent correct, and overconfidence for questions with and without treatment.

Confidence and percent correct are shown on the left vertical axis and overconfidence is shown

on the right vertical axis. Left Panel: Consider the Unknowns; Right Panel: Consider the

Alternative. Standard errors displayed.

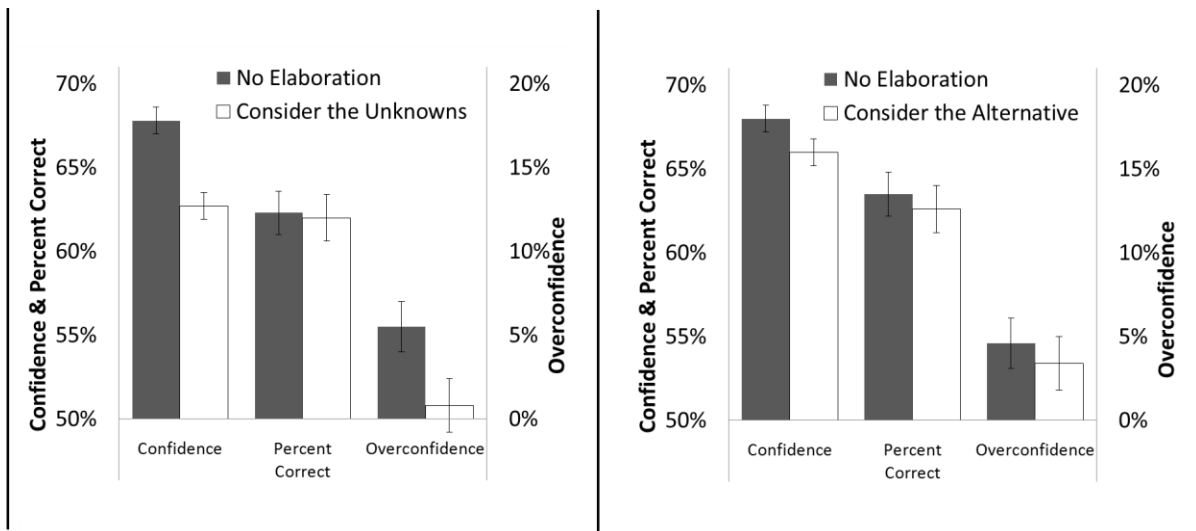


Figure 3: Effect of Consider the Unknowns Across Domains, Chapter 1, Study 3

Overconfidence on questions with and without treatment in overconfident domains and calibrated/underconfident domains. Left Panel: Consider the Unknowns; Right Panel: Consider the Alternative. Standard errors displayed.

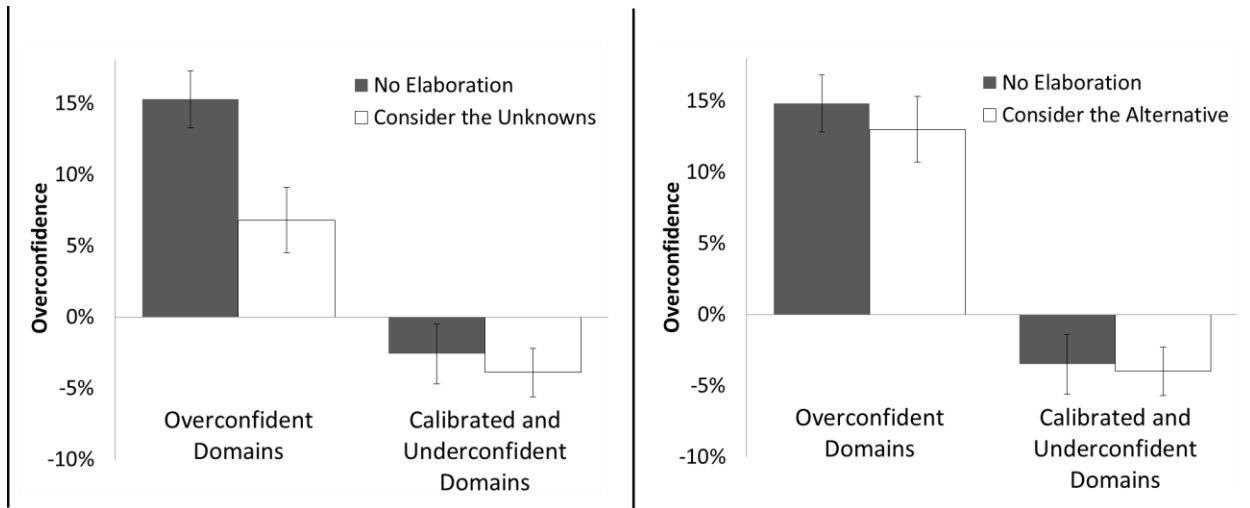


Figure 4: Effect of prime X financial Literacy Chapter 2, Study 2

The prime only influenced the asset allocation of those low in financial literacy. Asset concentration is shown on the vertical axis while financial literacy is shown on the horizontal axis. Financial literacy scores ranged from 4 to 11.

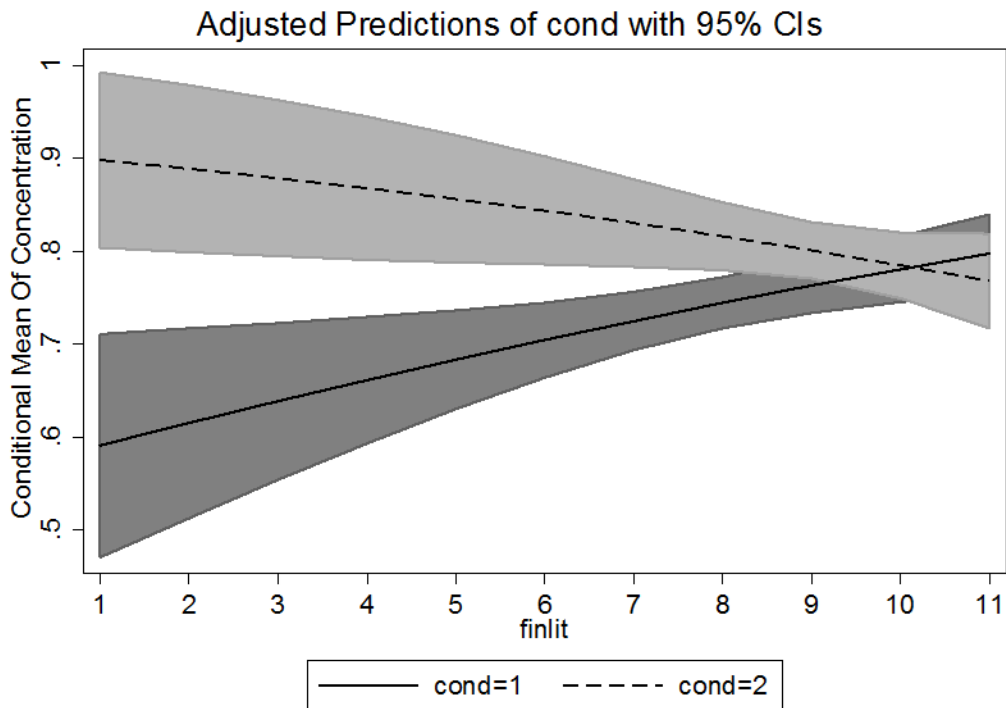
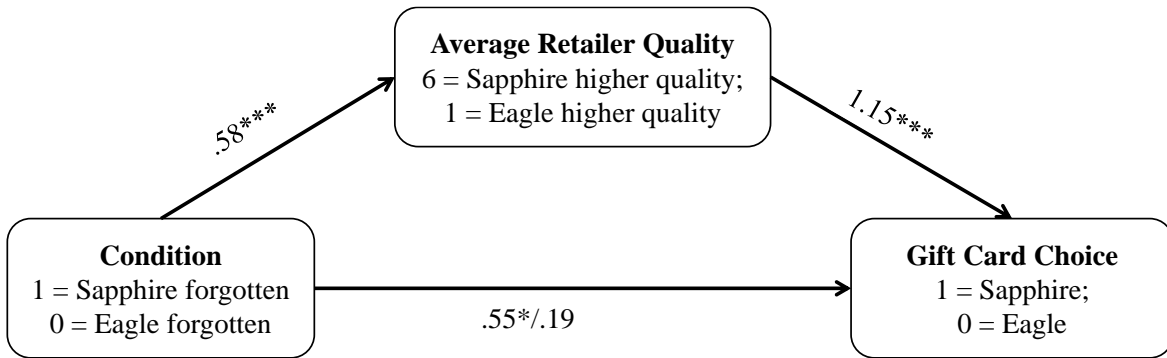


Figure 5: Mediation of Gift Card Choice, Chapter 3, Study 2

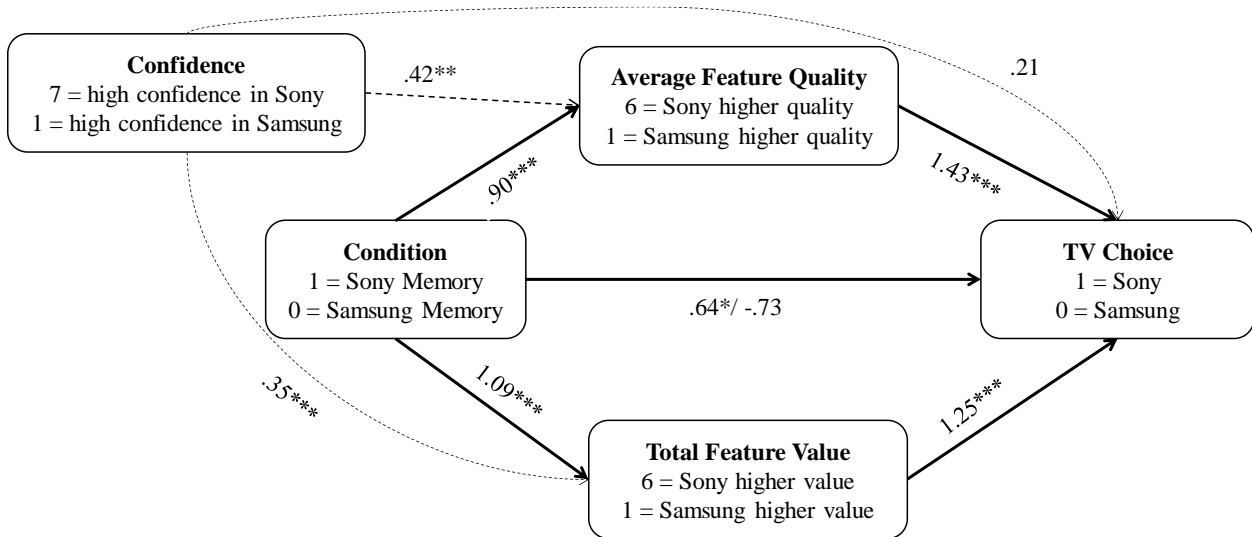
Average retailer quality fully mediates the effect of condition on choice.



* $p < .05$, ** $p < .01$, *** $p < .001$.

Figure 6: Mediation of TV Choice, Chapter 3, Study 4

Average feature quality and total feature value fully mediate the effect of condition on choice after controlling for confidence.



* $p < .05$, ** $p < .01$, *** $p < .001$.

Appendix A: Materials for Studies in Chapter 1

Study 1 Two Alternative Forced Choice Questions

<i>Question</i>	<i>Choice 1</i>	<i>Choice 2</i>
Q1: Which of these U.S. presidents held office first?	Andrew Jackson	Franklin Pierce
Q2: Which of these fast food items has more calories?	Subway meatball marinara foot-long sandwich	McDonald's double quarter pounder with cheese
Q3: Which of these states had a higher population in 2010?	Connecticut	Nevada
Q4: Which of these attractions had more visitors in 2007?	Great Wall of China	Empire State Building
Q5: Which of these 2011-model cars gets more miles per-gallon in real driving (mix of highway and city)?	Volkswagen Jetta	Audi A3 Quattro
Q6: Which of these "tourist cities" has a warmer daily high temperature in July, on average?	Rome, Italy	Sydney, Australia
Q7: Which of these U.S. universities charged higher tuition in 2010?	University of Chicago	Harvard University
Q8: Of these two "principal mountains of the world," which is taller?	Muztagh Ata, China	Mt. Elbrus, Russia

Q9: Which of these states had a higher percentage of its population with incomes below the federal poverty line in 2010?

Missouri

Delaware

Q10: Which of these cities is farther from Los Angeles, in air miles?

St. Louis

Denver

Study 1 Known unknown rating instructions

Instructions: Please read the following information carefully

To what degree were each of your reasons about **something that is unknown to you** versus **something that is known to you**?

Please use this as a guide to carefully rate each of your reasons. We will use the estimate '*Who is older Madonna or Celine Dion?*' as an example:

If your reason was about something that you are **sure you know**, such as "*Celine Dion released more records over her career than Madonna,*" then you might rate this a 6 or 7, since you are stating something **known to you**.

If your reason was about something that you **vaguely know**, such as "*I think Madonna has more children than Celine Dion,*" then you might rate this a 3 to 5 since you are stating something that is somewhere between **known and unknown to you**.

If your reason was about something **that you have no idea**, such as "*I have no idea if Madonna released her debut album before Celine Dion,*" then you might rate this a 1 or 2, since you are stating something **unknown to you**.

If your reason was **not very specific or useful**, but still communicated a fact about which you are certain, such as "*I know that Madonna is older than Britney Spears*" you will still rate this as a 7, since it is still something **known to you**.

This scale is **not** meant to measure how much each reason improved your estimate, or how relevant each reason was to your estimate. It is only meant to measure if the reason was about **something that is unknown to you** or **something that is known to you**.

As a reminder, You listed out reasons that you were certain or uncertain in your answer to:

Study 1, 2, and 3 Typical Text Responses

Study 1 reasons listed by participants for "Which of these U.S. presidents held office first? A. Andrew Jackson B. Franklin Pierce"

Statement rated as unknown	<i>"I do not know much about andrew jackson so i do not know how old of a president he is"</i>
Statement rated as known	<i>"Jackson is known as one of the best presidents, and the best ones came first in America's history."</i>
Statement coded in support of chosen answer for when choice is "B. Franklin Pierce"	<i>"Franklin Pierce seems like an old-fashioned name."</i>
Study 1 Statement in support of alternative answer when choice is "B. Franklin Pierce"	<i>"Andrew Jackson is on the twenty dollar bill."</i>

Study 2 statements by participants

Question: "Which of these countries has the highest life expectancy? A. Greece B. Finland C. Singapore D. United Kingdom"

Consider the Unknowns (CTU) statement	<i>"what are the economic conditions in Greece"</i>
Consider the Alternative (CTA) statement when choice is "B. Finland"	<i>"The correct answer might be singapore because it has high average income"</i>

Question: "Which movie made the most money at the box office in real dollars? A. The Empire Strikes Back B. Ben-Hur C. 101 Dalmatians D. The Godfather"

Consider the Unknowns (CTU) statement:	<i>I don't know how popular the theaters were when the movies came out</i>
Consider the Alternative (CTA) statement when choice is "A. Emprrie Strikes Back"	<i>"The Godfather is highly critically acclaimed, so it may have been seen in theaters more"</i>

Study 3 statements by participants

Question: ""Which of the following food items contains the least calories? A. Green beans (canned, drained, 1 cup) B. Chicken breast (boneless, skinless, roasted, 3 ounces)"

Consider the Unknowns (CTU) statement	<i>"What size can of green beans and is salt added?"</i>
---------------------------------------	--

Consider the Alternative (CTA)
statement when choice is "A. Green
beans"

*"Because the chicken breast
is both boneless and skinless,
the calorie count may be
lower than the green beans."*

***Question: "Which of these 2013 model cars gets the most miles per gallon in real
driving (mix of highway and city)? A. Lexus IS 350 AWD B. Nissan Pathfinder 4WD***

Consider the Unknowns (CTU)
statement:

*"I don't know the size of the engines
of these cars"*

Consider the Alternative (CTA)
statement when choice is "A. Lexus IS
350 AWD"

"It's possible the Nissan is a hybrid."

Study 2 Performance Incentive

Two participants will be randomly drawn to receive a bonus based on their performance in this task. If you are drawn you can receive a maximum total bonus of \$212 depending on (1) the number of questions you correctly answer and (2) the accuracy of your probability estimates. For each question correctly answered you will receive \$14. In addition, your assessed probability will be scored for accuracy by calculating a Brier Score. A Brier Score is a common scoring method for determining the accuracy or "calibration" of probabilistic predictions. More precisely, a Brier Score measures the mean squared difference between a person's probability judgments and the truth. Scores take on a value between zero and one, with lower Brier Scores reflecting greater accuracy and higher scores reflecting greater inaccuracy. The lower your Brier Score, the greater your bonus (A Brier Score of 0 will receive \$100 on top of the bonus for questions correct, and a brier score of .5 will receive a bonus of \$50, ect.). Thus, it is in your best interests to be as thoughtful and accurate as possible in all of your estimates. Please press 'enter' to begin the task.

Study 2 Questions

1. Which of these countries has the highest life expectancy? A. Greece B. Finland C. Singapore D. United Kingdom
2. Which movie made the most money at the box office in real dollars? A. The Empire Strikes Back B. Ben-Hur C. 101 Dalmatians D. The Godfather
3. Which state had the highest population as measured in July, 2012? A. Maryland B. Indiana C. Missouri D. Tennessee
4. Which element has the lowest atomic weight? A. Carbon B. Nitrogen C. Oxygen D. Fluorine
5. Which of the following food items contains the most calories? A. Ice cream (vanilla, 4 ounces) B. Chicken breast (boneless, skinless, roasted, 3 ounces) C. Ranch salad dressing (2 tablespoons) D. Hot dog (beef and pork)
6. Which city is the further from Kansas City, MO in air miles? A. Little Rock AR B. Denver, CO C. Cheyenne WY D. Minneapolis MN
7. Which of these beverages have the highest number of calories? A. Cranberry juice cocktail (12 ounces) B. Whole Milk (12 ounces) C. Beer (12 ounces) D. Hard liquor (vodka, rum, whiskey, gin; 80 proof) (1.5 ounces)
8. Which of these 2013 model cars gets the most miles per gallon in real driving (mix of highway and city)? A. Honda FIT B. Ford Taurus FWD C. Honda Accord D. Cadillac ATS

Study 3 question stems and number of questions.

Participants viewed 20 non repeated questions that were drawn from the following 9 databases of question domains. Participants answered an even proportion of questions from each domain on average.

1. Which of these countries has the highest life expectancy? (24,753 questions)
2. Which of the following food items contains the least calories? (44,253 questions)
3. Which city is the closest to Kansas City, MO in air miles? (1,711 questions)
4. Which President was elected first? (946 questions)
5. Which of these 2013 model cars gets the most miles per gallon in real driving (mix of highway and city)? (679,195 questions)
6. Which state had the highest population as measured in July, 2012? (1,225 questions)
7. Which of these beverages have the lowest number of calories? (300 questions)
8. Which movie made the most money at the box office in real dollars? (19,900 questions)
9. Which element has the lowest atomic weight? (5,886 questions)

Study 3 Procedure for identifying overconfident domains and calibrated/underconfident domains.

We used a split sample procedure to identify which domains people show overconfidence or calibrated/underconfident domains. The Study 3 sample can be split into 4 groups: (1) CTU condition, no treatment questions (2) CTU condition, treatment, (3) CTA condition, no treatment questions (4) CTA condition, treatment questions. We wished to test the difference in domain calibration between groups 1 and 2. Thus, we must establish a baseline level of domain calibration in a sample outside of these groups to avoid regression to the mean (Klayman et al., 1999). To accomplish this we used group 3 (CTA condition, no treatment questions) to identify for which domains participants exhibited overconfidence and for which domains participants exhibited underconfidence/calibration.

Appendix B: Materials for Studies in Chapter 2

Study 1-5, 6-item EARS Instructions:

The scale starts with a prompt about an evaluation or forecast. For example, if the forecast is evaluating the return of a stock the prompt would be:

Consider the task of *evaluating the approximate return of an individual stock over the course of 1 year*.

The approximate return of an individual stock over 1 year ...

The “approximate return of an individual stock” is then evaluated on epistemic and aleatory items as shown below.

The scale should appear as shown below with items 1-3 presented at the top of the scale (but randomized within the 1-3 group) and items 4-6 presented at the bottom of the scale (but randomized within the 4-6 group):

Consider the task of *{insert evaluation here}*.

{insert evaluation here} ...

	Not at all						Very much
... is something that has an element of randomness	o	o	o	o	o	o	o
... is determined by chance factors	o	o	o	o	o	o	o
... could play out in different ways on similar occasions	o	o	o	o	o	o	o
... is knowable in advance, given enough information	o	o	o	o	o	o	o
... is something that well-informed people would agree on	o	o	o	o	o	o	o
... is something that becomes more predictable with additional knowledge or skills	o	o	o	o	o	o	o

Scoring

The first three items are averaged to score perceptions of aleatory uncertainty and the last three items are averaged to score perception of epistemic uncertainty.

Study 1 and 2 Exclusion Criteria

Q115 Have you ever invested in the stock market? (i.e. purchased a single stock, or a stock mutual fund)?

- Yes (1)
- No (2)

If 2 is selected, then exclude

Q116 If yes, what is the current value of your stock market investments?

- \$0-\$100 (1)
- \$100-\$250 (2)
- \$250-\$500 (3)
- \$500-\$750 (4)
- \$750-\$1,000 (5)
- \$1,000-\$2,000 (6)
- \$2,000-\$5,000 (7)
- \$5,000-\$20,000 (8)
- \$20,000-\$50,000 (9)
- more than \$50,000 (10)

If 1-5 is selected, then exclude

Q132 To what extent do you rely on someone else to make investment decisions for you? (e.g. rely on a spouse or financial advisor to make decisions)

- I rely on someone else to make all of my investment decisions for me (1)
- I rely on someone else to make the vast majority of investment decisions for me (2)
- I rely on someone else to make most investment decisions for me (3)
- I rely on someone else to make some investment decisions for me (4)
- I rely on someone else to make a small minority of investment decisions for me (6)
- I make all of my investment decisions for myself (7)

If 1 or 2 selected, then exclude

Q129 How would you characterize your knowledge about investing in the stock market?

- I know nothing about investing in the stock market (1)
- I have a very small amount of knowledge about investing in the stock market (2)
- I know what investing in the stock market is but do not consider myself very knowledgeable in the subject (3)
- I know what investing in the stock market is and have a moderate level of knowledge in the subject (4)
- I consider myself an expert on stock market investing (5)

If 1 or 2 selected, then exclude

Q144 How well do you know what a stock is?

- I know nothing about what a stock is (1)
- I have a very small amount of knowledge about what a stock is (2)
- I know what a stock is but do not consider myself very knowledgeable in the subject (3)
- I know what a stock is and have a moderate level of knowledge in the subject (4)
- I know what a stock is and consider myself an expert in the subject (5)

If 1 or 2 selected, then exclude

Q146 How well do you know what a mutual fund is?

- I know nothing about what a mutual fund is (1)
- I have a very small amount of knowledge about what a mutual fund is (2)
- I know what a mutual fund is but do not consider myself very knowledgeable in the subject (3)
- I know what a mutual fund is and have a moderate level of knowledge in the subject (4)
- I know what a mutual fund is and consider myself an expert in the subject (5)

If 1 or 2 selected, then exclude

Q143 How well do you know what an index fund is?

- I know nothing about what an index fund is (1)
- I have a very small amount of knowledge about what an index fund is (2)
- I know what an index fund is but do not consider myself very knowledgeable in the subject (3)
- I know what an index fund is and have a moderate level of knowledge in the subject (4)
- I know what an index fund is and consider myself an expert in the subject (5)

If 1 or 2 selected, then exclude

Q121 An index fund is:

- A form of debt (1)
- Is actively managed (2)
- Is usually not diversified (3)
- Is often composed of many assets (4)
- I don't know (5)

If 4 is not selected, then exclude

Q122 A mutual fund is:

- A type of insurance (1)
- Never actively managed (2)
- Is most commonly used to purchase real estate (3)
- An asset that can be composed of stocks or bonds (4)
- I don't know (5)

If 4 is not selected, then exclude

Q124 A stock is:

- Something that's only issued by technology companies (1)
- Something that's only purchased by investment bankers (2)
- A share of a company (3)
- A type of insurance (4)
- I don't know (5)

If 3 is not selected, then exclude

AF What is one primary purpose of the stock market? Please select "Other" and type in the word "survey" in the box provided rather than answering this question because it is an attention check.

- Allow people to trade stocks (1)
- Price bonds (2)
- Fund public schools (3)
- Prevent natural disasters (4)
- Other (5) _____

If 5 is not selected, then exclude

Study 1 DOSPERT questions

People often see some risk in situations that contain uncertainty about what the outcome or consequences will be and for which there is the possibility of negative consequences. However, riskiness is a very personal and intuitive notion, and we are interested in your gut level assessment of how risky each situation or behavior is. For each of the following statements, please indicate how risky you perceive each situation. Provide a rating from Not at all Risky to Extremely Risky, using the following scale:

	Not at all Risky 1 (1)	Slightly Risky 2 (2)	Somewhat Risky 3 (3)	Moderately Risky 4 (4)	Risky 5 (5)	Very Risky 6 (6)	Extremely Risky 7 (7)
Investing 10% of your annual income in a moderate growth mutual fund. (1)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Investing 5% of your annual income in a very speculative stock. (2)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Investing 10% of your annual income in a new business venture. (3)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Study 3 Primes:

High Epistemic Prime:

Greatest Investor of All Time

Peter Lynch's spot in the list of greatest investors stems in large part from his work with the Fidelity Magellan Fund between 1977 and 1990. During this 13-year period, the fund posted an annual average return of 29%, beating the S&P 500 index in 11 out of 13 years and building the fund's assets from \$20 million to \$14 billion. The fund had the best 20-year return rate of any mutual fund in history.

Investment Philosophy:

Peter Lynch made use of a set of eight fundamental principles when selecting stocks to purchase. He has distributed that checklist to younger investors and the public at various conferences and talks:

- Know what you own.
- You have plenty of time to identify and recognize exceptional companies.
- Avoid long shots.
- Good management is very important - buy good businesses.
- Be flexible and humble, and learn from mistakes.
- Before you make a purchase, you should be able to explain why you're buying.
- There's always something to worry about.

The first of these principles was especially important to Lynch. He always stuck with companies and industries that he knew about or could easily learn about. Along with that principle was a steadfast commitment to due diligence before making any investment decision. Lynch was famous for focusing on a company's fundamentals, using a bottom-up approach. He only invested for the long run and paid little attention to short-term market fluctuations.

Quotes:

"Go for a business that any idiot can run – because sooner or later, any idiot is probably going to run it."
"If you stay half-alert, you can pick the spectacular performers right from your place of business or out of the neighborhood shopping mall, and long before Wall Street discovers them."
"Investing without research is like playing stud poker and never looking at the cards."

Low Epistemic Prime:

It's Official! Gurus Can't Accurately Predict Markets

CXO Advisory Group has been collecting data from market forecasters since 1998. The firm has tracked and graded thousands of market forecasts made by dozens of popular gurus over the years. The overall results are not good. CXO has concluded that the market experts accurately predicted market direction less than half the time.

Although gurus tended to come and go during their collection period, CXO found forecasting ability tends to stay about the same.

It only took 2 years and about 200 predictions before the accuracy rating fell below half the time in early 2000. The cumulative accuracy has stayed at that low level half the ever since. By 2008, CXO had collected and graded more than 5,000 predictions.

The consistent low accuracy grade has led CXO to make a command decision. They will stop collecting and grading new forecasts in 2013. The preliminary study on the CXO website graded 6,459 forecasts. Their final report will include an additional 126 grades for 2012, although they are not expected to materially affect the preliminary results.

I am sad to see CXO end their guru grade tracking. It provides an excellent real time example of how sales skill triumphs over investment skill on Wall Street. Forecasting isn't about predicting the market; it's about marketing the prediction. As one newsletter guru told me years ago, "Given a choice between great marketing and great forecasting, I'd pick great marketing every time."

Study 3 Stock Information

STOCK PRICE RESEARCH REPORT

Aon PLC (Ticker: AON)

The top 25 stock analysts project the stock price of Aon will **decrease by 12.3%** over the next 6 month.

Nabors Industries (Ticker: NBR)

The top 25 stock analysts project the stock price of Nabors will **increase by 21.6%** over the next 6 month.

Textron Inc (Ticker: TXT)

The top 25 stock analysts project the stock price of Textron will **decrease by 9.5%** over the next 6 month.

Disclaimer: Please do not distribute these stock estimates as they are only licensed for use in this study.

Study 4 Research Report



Global Research

28 February 2017

Apple Inc. How Augmented Reality Could be the Next Big Thing for Apple

Equities	
Americas	
Computers	
12-month rating	Buy
12m price target	US\$151.00

AR will take time but could be a boost to Apple

Augmented reality (AR) is the use of computing to digitally enhance the real world. Apple appears more interested in AR, which connects people, than VR, which potentially isolates experiences. Our work suggests that AR could be the next major innovation from Apple and that its competencies could make the company a winner. Industry sources are upbeat and anxious to see what Apple does. Investors could be surprised at how AR could reinvigorate the iPhone/iPad and possibly result in new products. We raise our price target to \$151 based on improved upgrade and retention rates.

Cook compares AR to iPhone: "I think AR is that big, it's huge"

CEO Tim Cook has commented on Apple's interest in augmented reality, suggesting that AR will be huge and is a technology and not a product per se but also cautioning that while profound, AR will take some time to get right. As one developer told us, "Looking down at the screen and closed off in our own world will seem like the dark ages." Thanks to advanced cameras, consumers will hold their phones up with images superimposed onto the screen in cars, rooms, or walking down the street. 3D mapping through Simultaneous Localization and Mapping (SLAM) will be key.

Early AR might come with iPhone 8

Participants in AR include Google Tango, Magic Leap, and Microsoft HoloLens, which has been pushed out to 2019. Apple tends not to be first to market but often wins first to mind. Its advantages could include: (1) hardware expertise and superior hardware/software integration; (2) consistent updated releases with most customers on the latest version of iOS compared with the fragmented Android OS base; (3) an installed base of iPhone and iPad customers that can use AR rather than starting from scratch; and (4) a cloud infrastructure that facilitates data gathering. We layout an introduction timeline and expect the next iPhone could include moderate 3D mapping using stereoscopic vision and possibly an AR software development kit.

Valuation: Increase target price to \$151

We are increasing our price target from \$138 to \$151 based on a P/E of 16.5x F17e of \$9.10 (vs 15x previously), derived from 5-8% NOPAT growth, up from 3-6%, over the next eight years at a 20% or greater ROIC ([link](#)). Ex-cash the target P/E is 14x.

Steven Milunovich, CFA

Analyst
steven.milunovich@ubs.com
+1-212-713 3372

Benjamin Wilson, CPA
Associate Analyst
benjamin-d.wilson@ubs.com
+1-212-713 1442

Appendix C: Materials for Studies in Chapter 3

Study 1 Retailers

Hibbett Sports	J.C. Penney Co.
Aqua Flow	Le Gourmet Chef
Family Dollar Stores	Foot Locker
Hobby Lobby	Burger King Worldwide
Hallmark Business Expressions	Williams-Sonoma
Newbury Comics	Build-A-Bear Workshop
Annas Linens	Disney Store
Party City	Sams Club
Sonic	Green Earth Market
All Pro Sound	Jo-Ann Stores
Payless ShoeSource	Dominos Pizza
Festival Foods (Minnesota)	Paper Source
Claire's	Anthropologie
REI	Bed Bath & Beyond
Lucy Activewear	Petco
Charming Charlie	Chipotle Mexican Grill
HomeGoods	Free People
Hertz Car Sales	Michaels
Price Chopper Supermarkets	McDonalds
Hobby Express	7-Eleven
Kirklands	Office Depot
Famous Footwear	Whole Foods Market
Neiman Marcus	iTunes
PetSmart	Ikea North America Services
OfficeMax	Bath & Body Works
Crate & Barrel	Starbucks
BevMo!	Apple Inc.
Subway	Nordstrom
Panera Bread Company	Costco
Pottery Barn	Trader Joes

Study 2 Retailers

Panera Bread Company	Big Lots
Ikea North America Services	J.C. Penney Co.
Apple Inc.	Sherwin-Williams
Dell	Kirkland's
The Home Depot	HomeGoods
Disney Store	Famous Footwear
Bath & Body Works	Staples
7-Eleven	Michaels Stores
Whole Foods Market	HobbyTown USA
Trader Joe's	Le Gourmet Chef
PetSmart	Tempur-Pedic
Pottery Barn	Hallmark Business Expressions
Subway	REI
Costco	L.L.Bean
Barnes & Noble	Build-A-Bear Workshop
Foot Locker	Gap
Sam's Club	KitSplit
Sonic	Michael C. Fina
Dick's Sporting Goods	Full Compass Systems
Petco	CST Brands
Family Dollar Stores	Hancock Fabrics
Toys 'R' Us	Defense Commissary Agency
Michaels	Adorama
Starbucks	Raymour & Flanigan
Bed Bath & Beyond	Lauren Moshi
Dunkin' Donuts	Relax The Back
Nordstrom	Signet Jewelers
Dollar Tree	Rainbow Shops
Dollar General	Crystal Rock Holdings
Crate & Barrel	Festival Foods (Minnesota)
Party City	Alfred Angelo
Wendy's	Fleming Companies, Inc
Office Depot	J. Press
Claire's	Party America
AT&T Wireless	North Shore Shrimp Trucks
Chipotle Mexican Grill	Mor Furniture
Burger King Worldwide	Bluemercury
Ace Hardware	Mecox
Cracker Barrel	Lithia Motors
Williams-Sonoma	Deseret Industries
BJ's Wholesale Club	Orion Telescopes & Binoculars
Chick-fil-A	Battery Giant
McDonald's	CardCash
Kitchen Collection	Davenport Cash Store

Study 3 Movies

The Godfather	Apocalypse Now
The Shawshank Redemption	Amadeus
Schindler's List	The Lord of the Rings: The Return of the King
Raging Bull	Gladiator
Casablanca	Titanic
Citizen Kane	From Here to Eternity
Gone with the Wind	Saving Private Ryan
The Wizard of Oz	Unforgiven
One Flew Over the Cuckoo's Nest	Raiders of the Lost Ark
Lawrence of Arabia	Rocky
Vertigo	A Streetcar Named Desire
Psycho	The Philadelphia Story
The Godfather: Part II	To Kill a Mockingbird
On the Waterfront	An American in Paris
Sunset Blvd.	The Best Years of Our Lives
Forrest Gump	My Fair Lady
The Sound of Music	A Clockwork Orange
12 Angry Men	Doctor Zhivago
West Side Story	The Searchers
Star Wars: Episode IV - A New Hope	Jaws
2001: A Space Odyssey	Patton
E.T. the Extra-Terrestrial	Butch Cassidy and the Sundance Kid
The Silence of the Lambs	The Treasure of the Sierra Madre
Chinatown	The Good, the Bad and the Ugly
The Bridge on the River Kwai	The Apartment
Singin' in the Rain	Platoon
It's a Wonderful Life	High Noon
Dr. Strangelove or: How I Learned to Stop Worrying and Love the Bomb	Braveheart
Some Like It Hot	Dances with Wolves
Ben-Hur	Jurassic Park

Study 4 TV Features

LCD Screen
Harmon Kardon speaker system
Touch screen remote
RokuTV included
Weight: 40 lbs
Power consumption: 80 watts
Blue-ray player included
1 year warranty included
Multi room link to connect sound systems across rooms
Dolby digital audio
Bluetooth headset support
DVD player
Audio depth enhancing technology
Surround sound ready
LED back lighting
Quadcore processor
DirectTV ready
Multi lingual display
Eco sensor to adapt the screen's brightness to the intensity of the light in the room
Connectshare to watch videos, play music, or view photos through a USB connection
Ultraclear panel technology that brightens colors and enhances contrast
RokuTV included
Curved TV
3D TV technology
HDMI Port
Quantum Dots technology
Voice recognition remote control
240HZ refresh rate
AppleTV included
1 year Netflix subscription included
1 year Amazon Prime subscription included
1 year Hulu subscription included
1 year HBO subscription included
1 year Seeso subscription included
1 year Showtime subscription included
Motion flow technology
V-Chip to allow parents to block inappropriate movies and programs
Apps platform to connect all of your favorite apps
SmartView to broadcast your TV to your smartphone
Clik Connect to broadcast your smartphone to your TV
Closed captioning
One-connect box for connect all of your devices
USB port
Ethernet port
Optical digital audio output
Coaxial cable port
Energystar certified
Wallmount ready
Removable stand
Internal TV antenna

Study 5 Negative Reviews

It could use more HDMI ports
It does not have internet connectivity
After 2 months, the TV started make a loud buzz every 45 minutes
Closed captioning seems slightly delayed
Good (not very good or great) contrast ratio. Perfectly acceptable, however. "Blacks" are black and bright colors vivid against them
I had to call a technician to set up DirectTV
I have a sunny living room and there is quite a bit of glare
Image is okay, nothing to write home about.
Inside of screen was found to be cracked so there is a long scratch defect appearing when turned on.
It didn't work. I sent it back.
Lacks advanced picture settings of higher spec sets
Netflix loads slowly on this TV compared to my laptop
Only has spanish and english language options
Remote instructions are confusing
Screen starts to washout some at around a 30 degree angle
Seems to consume a lot of power
Short warranty of only 6 months
sometimes does not play well with Roku
the mounting threads for your wall mount are not a very standard size
The picture quality is ok for HD but really poor for non-HD channels
The placement of the hdmi and other inputs is inconvenient if you plan to use a wall mount
The sound definition isn't clear
The stand is too wide for my table
The TV seems unusually heavy
The tv works fine with my Roku and antenna, but not with my mac mini
there are only 2 HDMI ports
There is no DVD player on the TV
There is no ethernet port
There is no internal BlueRay
This TV does not come with composite Hookups
TV came with a plasma screen, I prefer LCD
Tv doesn't have an optical digital audio output
tv stand seems a bit flimsy
TV turns on slowly
TV was damaged during delivery
TV won't stay turned off, I have to unplug it
Viewing angles are adequate
When I set up and plugged in the set the screen was cracked and the unit would not function
Won't connect to my bluetooth
You must purchase a wallmount separately
Plastic casing seems flimsy

ESPN is too dark

TV suffers small flickering and judder during faster motion

the sections light don't blend together so it looks like multiple rectangles on your screen

The network port failed after 3 days

Does not pick up wifi signal. Extremely unsatisfied.

Does not support pandora

Apps do not include spotify

The delivery took forever to arrive (way past promised date!)

It was very difficult to remove this TV from the cardboard box

References

- Alba, Joseph W., Howard Marmorstein, and Amitava Chattopadhyay (1992), "Transitions in preference over time: The effects of memory on message persuasiveness," *Journal of Marketing Research*, 406–416.
- Alpert, Marc and Howard Raiffa (1982), "A progress report on the training of probability assessors."
- Alter, Adam L., Daniel M. Oppenheimer, and Jeffrey C. Zemla (2010), "Missing the trees for the forest: A construal level account of the illusion of explanatory depth.," *Journal of Personality and Social Psychology*, 99 (3), 436.
- Babad, Elisha (1987), "Wishful thinking and objectivity among sports fans.," *Social Behaviour*.
- Basu, Sanjoy (1977), "Investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient market hypothesis," *The journal of Finance*, 32 (3), 663–682.
- Baumann, Andrea O., Raisa B. Deber, and Gail G. Thompson (1991), "Overconfidence among physicians and nurses: the 'micro-certainty, macro-uncertainty' phenomenon," *Social science & medicine*, 32 (2), 167–174.
- Beckwith, Neil E. and Donald R. Lehmann (1975), "The importance of halo effects in multi-attribute attitude models," *Journal of Marketing research*, 265–275.
- Ben-David, Itzhak, John R. Graham, and Campbell R. Harvey (2007), "Managerial overconfidence and corporate policies," National Bureau of Economic Research.
- Bettman, James R. (1979), "Memory factors in consumer choice: A review," *The Journal of marketing*, 37–53.

- Biehal, Gabriel and Dipankar Chakravarti (1982), "Information-presentation format and learning goals as determinants of consumers' memory retrieval and choice processes," *Journal of Consumer Research*, 8 (4), 431–441.
- and ——— (1983), "Information accessibility as a moderator of consumer choice," *Journal of Consumer Research*, 10 (1), 1–14.
- and ——— (1986), "Consumers' use of memory and external information in choice: Macro and micro perspectives," *Journal of Consumer Research*, 12 (4), 382–405.
- Bothwell, Robert K., Kenneth A. Deffenbacher, and John C. Brigham (1987), "Correlation of eyewitness accuracy and confidence: Optimality hypothesis revisited.," *Journal of Applied Psychology*, 72 (4), 691.
- Bradlow, Eric T., Ye Hu, and Teck-Hua Ho (2004), "A learning-based model for imputing missing levels in partial conjoint profiles," *Journal of Marketing Research*, 41 (4), 369–381.
- Brenner, Lyle, Dale Griffin, and Derek J. Koehler (2005), "Modeling patterns of probability calibration with random support theory: Diagnosing case-based judgment," *Organizational Behavior and Human Decision Processes*, 97 (1), 64–81.
- Broniarczyk, Susan M. and Joseph W. Alba (1994), "The role of consumers' intuitions in inference making," *Journal of Consumer Research*, 21 (3), 393–407.
- Brown, Alexander L., Colin F. Camerer, and Dan Lovallo (2012), "To review or not to review? Limited strategic thinking at the movie box office," *American Economic Journal: Microeconomics*, 4 (2), 1–26.
- Brown, Roger and James Kulik (1977), "Flashbulb memories," *Cognition*, 5 (1), 73–99.

- Buehler, Roger, Dale Griffin, and Michael Ross (1994), "Exploring the " planning fallacy": Why people underestimate their task completion times.," *Journal of personality and social psychology*, 67 (3), 366.
- Burke, Alafair, Friderike Heuer, and Daniel Reisberg (1992), "Remembering emotional events," *Memory & cognition*, 20 (3), 277–290.
- Burke, Raymond R. and Thomas K. Srull (1988), "Competitive interference and consumer memory for advertising," *Journal of consumer research*, 15 (1), 55–68.
- Camerer, Colin and Dan Lovallo (1999), "Overconfidence and excess entry: An experimental approach," *The American Economic Review*, 89 (1), 306–318.
- Caputo, Deanna and David Dunning (2005), "What you don't know: The role played by errors of omission in imperfect self-assessments," *Journal of Experimental Social Psychology*, 41 (5), 488–505.
- Chernev, Alexander and Gregory S. Carpenter (2001), "The role of market efficiency intuitions in consumer choice: A case of compensatory inferences," *Journal of Marketing Research*, 38 (3), 349–361.
- Christianson, Sven-Åke (1992), "Emotional stress and eyewitness memory: a critical review.," *Psychological bulletin*, 112 (2), 284.
- and Lars-Göran Nilsson (1984), "Functional amnesia as induced by a psychological trauma," *Memory & Cognition*, 12 (2), 142–155.
- Clifford, Brian R. and Clive R. Hollin (1981), "Effects of the type of incident and the number of perpetrators on eyewitness memory.," *Journal of Applied Psychology*, 66 (3), 364.
- and Jane Scott (1978), "Individual and situational factors in eyewitness testimony.," *Journal of Applied Psychology*, 63 (3), 352.

- Cooper, William H. (1981), "Ubiquitous halo.," *Psychological bulletin*, 90 (2), 218.
- Craik, Fergus IM, Richard Govoni, Moshe Naveh-Benjamin, and Nicole D. Anderson (1996), "The effects of divided attention on encoding and retrieval processes in human memory.," *Journal of Experimental Psychology: General*, 125 (2), 159.
- and Robert S. Lockhart (1972), "Levels of processing: A framework for memory research," *Journal of verbal learning and verbal behavior*, 11 (6), 671–684.
- Cutts, Martin (2013), *Oxford guide to plain English*, OUP Oxford.
- Dick, Alan, Dipankar Chakravarti, and Gabriel Biehal (1990), "Memory-based inferences during consumer choice," *Journal of Consumer Research*, 17 (1), 82–93.
- Downing, Cathryn J., Robert J. Sternberg, and Brian H. Ross (1985), "Multicausal inference: Evaluation of evidence in causally complex situations.," *Journal of Experimental Psychology: General*, 114 (2), 239.
- Dutta, Satrajit, Rabindra N. Kanungo, and Vaira Freibergs (1972), "Retention of affective material: Effects of intensity of affect on retrieval.," *Journal of Personality and Social Psychology*, 23 (1), 64.
- Ellsberg, Daniel (1961), "Risk, ambiguity, and the Savage axioms," *The quarterly journal of economics*, 643–669.
- Erev, Ido, Thomas S. Wallsten, and David V. Budescu (1994), "Simultaneous over-and underconfidence: The role of error in judgment processes.," *Psychological review*, 101 (3), 519.
- Ericson, Keith M. Marzilli (2011), "Forgetting we forget: Overconfidence and memory," *Journal of the European Economic Association*, 9 (1), 43–60.

- Feiler, Daniel C., Jordan D. Tong, and Richard P. Larrick (2013), "Biased judgment in censored environments," *Management Science*, 59 (3), 573–591.
- Fernbach, Philip M., Todd Rogers, Craig R. Fox, and Steven A. Sloman (2013), "Political extremism is supported by an illusion of understanding," *Psychological science*, 24 (6), 939–46.
- Ferrell, William R. and Paul J. McGoey (1980), "A model of calibration for subjective probabilities," *Organizational Behavior and Human Performance*, 26 (1), 32–53.
- Fischhoff, Baruch, Paul Slovic, and Sarah Lichtenstein (1978), "Fault trees: Sensitivity of estimated failure probabilities to problem representation.," *Journal of Experimental Psychology: Human Perception and Performance*, 4 (2), 330.
- Ford, Gary T. and Ruth Ann Smith (1987), "Inferential beliefs in consumer evaluations: An assessment of alternative processing strategies," *Journal of Consumer Research*, 14 (3), 363–371.
- Fox, Craig R. (1999), "Strength of evidence, judged probability, and choice under uncertainty," *Cognitive Psychology*, 38 (1), 167–189.
- and Robert T. Clemen (2005), "Subjective probability assessment in decision analysis: Partition dependence and bias toward the ignorance prior," *Management Science*, 51 (9), 1417–1432.
- and Gülden Ülkümen (2011), "Distinguishing two dimensions of uncertainty," *Perspectives on thinking, judging, and decision making*, 1–14.
- Gigerenzer, Gerd, Ulrich Hoffrage, and Heinz Kleinbölting (1991), "Probabilistic mental models: A Brunswikian theory of confidence.," *Psychological review*, 98 (4), 506.

- Giglio, Stefano and Kelly Shue (2014), “No News is News: Do Markets Underreact to Nothing?,” *Review of Financial Studies*, 27 (12), 3389–3440.
- Gilovich, Thomas, Robert Vallone, and Amos Tversky (1985), “The hot hand in basketball: On the misperception of random sequences,” *Cognitive psychology*, 17 (3), 295–314.
- Gneezy, Uri, John A. List, and George Wu (2006), “The uncertainty effect: When a risky prospect is valued less than its worst possible outcome,” *The Quarterly Journal of Economics*, 121 (4), 1283–1309.
- Gordon, Michael R. and Bernard E. Trainor (2006), *Cobra II: The inside story of the invasion and occupation of Iraq*, Vintage.
- Griffin, Dale and Lyle Brenner (2004), “Perspectives on probability judgment calibration,” *Blackwell handbook of judgment and decision making*, 177–199.
- and Amos Tversky (1992), “The weighing of evidence and the determinants of confidence,” *Cognitive psychology*, 24 (3), 411–435.
- Gunasti, Kunter and William T. Ross (2009), “How inferences about missing attributes decrease the tendency to defer choice and increase purchase probability,” *Journal of Consumer Research*, 35 (5), 823–837.
- Guynn, Melissa J., Mark A. McDaniel, and Gilles O. Einstein (1998), “Prospective memory: When reminders fail,” *Memory & Cognition*, 26 (2), 287–298.
- Hacking, Ian (1975), “The Emergence of Probability: A Philosophical Study of Early Ideas about Probability, Induction and Statistical Inference: A Philosophical Study of Early... Induction and Statistical Inference.”

- Hamill, Ruth, Timothy D. Wilson, and Richard E. Nisbett (1980), "Insensitivity to sample bias: Generalizing from atypical cases.," *Journal of Personality and Social Psychology*, 39 (4), 578.
- Hastie, Reid and Purohit A. Kumar (1979), "Person memory: Personality traits as organizing principles in memory for behaviors.," *Journal of Personality and Social Psychology*, 37 (1), 25.
- Heuer, Friderike and Daniel Reisberg (1990), "Vivid memories of emotional events: The accuracy of remembered minutiae," *Memory & Cognition*, 18 (5), 496–506.
- Higgins, E. Tory (1998), "Promotion and prevention: Regulatory focus as a motivational principle," *Advances in experimental social psychology*, 30, 1–46.
- Hoch, Stephen J. (1984), "Availability and interference in predictive judgment.," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10 (4), 649.
- (1985), "Counterfactual reasoning and accuracy in predicting personal events.," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11 (4), 719.
- Hong, Claire Yurong and Frank Weikai Li (2015), "The Information Content of Insider Sudden Silence," *Available at SSRN 2701254*.
- Howell, William C. and Sarah A. Burnett (1978), "Uncertainty measurement: A cognitive taxonomy," *Organizational Behavior and Human Performance*, 22 (1), 45–68.
- Huber, Joel and John McCann (1982), "The impact of inferential beliefs on product evaluations," *Journal of Marketing Research*, 324–333.
- Iyengar, Sheena S. and Mark R. Lepper (2000), "When choice is demotivating: Can one desire too much of a good thing?," *Journal of personality and social psychology*, 79 (6), 995.

- Jacoby, Jacob, Donald E. Speller, and Carol A. Kohn (1974), "Brand choice behavior as a function of information load," *Journal of Marketing Research*, 63–69.
- John, Deborah Roedder, Carol A. Scott, and James R. Bettman (1986), "Sampling data for covariation assessment: The effect of prior beliefs on search patterns," *Journal of Consumer Research*, 13 (1), 38–47.
- Johnson, D. D. (2004), *Overconfidence and war*, Cambridge: Harvard University Press.
- Johnson, Eric J., Gerald Häubl, and Anat Keinan (2007), "Aspects of endowment: a query theory of value construction.," *Journal of experimental psychology: Learning, memory, and cognition*, 33 (3), 461.
- Kahneman, D. and A. Tversky (1979), "Intuitive prediction: Biases and corrective procedures," *Management Science*, 12, 313–27.
- Kahneman, Daniel (2011), *Thinking, fast and slow*, Macmillan.
- , Barbara L. Fredrickson, Charles A. Schreiber, and Donald A. Redelmeier (1993), "When more pain is preferred to less: Adding a better end," *Psychological science*, 4 (6), 401–405.
- and Amos Tversky (1973), "On the psychology of prediction.," *Psychological review*, 80 (4), 237.
- and ——— (1982), "Variants of uncertainty," *Cognition*, 11 (2), 143–157.
- Kardes, Frank R. and David M. Sanbonmatsu (1993), "Direction of comparison, expected feature correlation, and the set-size effect in preference judgment," *Journal of Consumer Psychology*, 2 (1), 39–54.
- Keller, Kevin Lane (1987), "Memory factors in advertising: The effect of advertising retrieval cues on brand evaluations," *Journal of Consumer Research*, 14 (3), 316–333.

- , Susan E. Heckler, and Michael J. Houston (1998), “The effects of brand name suggestiveness on advertising recall,” *The Journal of Marketing*, 48–57.
- Kent, Robert J. and Chris T. Allen (1994), “Competitive interference effects in consumer memory for advertising: the role of brand familiarity,” *The Journal of Marketing*, 97–105.
- Kivetz, Ran and Itamar Simonson (2000), “The effects of incomplete information on consumer choice,” *Journal of Marketing Research*, 37 (4), 427–448.
- Klayman, Joshua (1995), “Varieties of confirmation bias,” *Psychology of learning and motivation*, 32, 385–418.
- and Young-Won Ha (1987), “Confirmation, disconfirmation, and information in hypothesis testing.,” *Psychological review*, 94 (2), 211.
- , Jack B. Soll, Claudia González-Vallejo, and Sema Barlas (1999), “Overconfidence: It depends on how, what, and whom you ask,” *Organizational behavior and human decision processes*, 79 (3), 216–247.
- Koehler, Derek J., Lyle Brenner, and Dale Griffin (2002), “The Calibration of Expert Judgment: Heuristics and Biases Beyond the Laboratory,” in *The calibration of expert judgment: Heuristics and biases beyond the laboratory*. In T. Gilovich, D. Griffin, & D. Kahneman (2002) *Heuristics and biases: The psychology of intuitive judgment*, Cambridge, MA: Cambridge University Press., 686–715.
- Koriat, Asher (2012), “The self-consistency model of subjective confidence.,” *Psychological review*, 119 (1), 80.
- , Sarah Lichtenstein, and Baruch Fischhoff (1980), “Reasons for confidence.,” *Journal of Experimental Psychology: Human learning and memory*, 6 (2), 107.

- Kunda, Ziva (1990), "The case for motivated reasoning.," *Psychological bulletin*, 108 (3), 480.
- Larrick, Richard P. (1993), "Motivational factors in decision theories: The role of self-protection.," *Psychological Bulletin*, 113 (3), 440.
- Lee, Angela Y. (2002), "Effects of implicit memory on memory-based versus stimulus-based brand choice," *Journal of Marketing Research*, 39 (4), 440–454.
- Lee, Byung-Kwan and Wei-Na Lee (2004), "The effect of information overload on consumer choice quality in an on-line environment," *Psychology & Marketing*, 21 (3), 159–183.
- Lewis, Donald J. (1979), "Psychobiology of active and inactive memory," *Psychological bulletin*, 86 (5), 1054.
- Li, Cong (2010), "Primacy effect or recency effect? A long-term memory test of Super Bowl commercials," *Journal of Consumer Behaviour*, 9 (1), 32–44.
- Lockwood, Penelope, Christian H. Jordan, and Ziva Kunda (2002), "Motivation by positive or negative role models: regulatory focus determines who will best inspire us," *Journal of personality and social psychology*, 83 (4), 854.
- Loftus, Elizabeth F. and Terrence E. Burns (1982), "Mental shock can produce retrograde amnesia," *Memory & Cognition*, 10 (4), 318–323.
- Lusardi, Annamaria, Olivia S. Mitchell, and Vilsa Curto (2010), "Financial literacy among the young," *Journal of consumer affairs*, 44 (2), 358–380.
- Lynch, John G. and Thomas K. Srull (1982), "Memory and attentional factors in consumer choice: Concepts and research methods," *Journal of Consumer Research*, 9 (1), 18–37.
- Mahajan, Jayashree (1992), "The overconfidence effect in marketing management predictions," *Journal of Marketing Research*, 29 (3), 329.

- Malhotra, Naresh K. (1982), "Information load and consumer decision making," *Journal of consumer research*, 8 (4), 419–430.
- Malmendier, Ulrike and Geoffrey Tate (2005), "CEO overconfidence and corporate investment," *The journal of finance*, 60 (6), 2661–2700.
- Markowitz, Harry (1952), "Portfolio selection," *The journal of finance*, 7 (1), 77–91.
- Matt, Georg E., Carmelo Vázquez, and W. Keith Campbell (1992), "Mood-congruent recall of affectively toned stimuli: A meta-analytic review," *Clinical Psychology Review*, 12 (2), 227–255.
- McClelland, Alastair GR and Fergus Bolger (1994), "The calibration of subjective probability: Theories and models 1980–94.," in *The calibration of subjective probability: Theories and models 1980–94.*, Oxford, England: John Wiley & Sons, 453–82.
- Meyer, Robert J. (1981), "A model of multiattribute judgments under attribute uncertainty and informational constraint," *Journal of Marketing Research*, 428–441.
- Moore, Don A. and Paul J. Healy (2008), "The trouble with overconfidence.," *Psychological review*, 115 (2), 502.
- Morewedge, Carey K., Daniel T. Gilbert, and Timothy D. Wilson (2005), "The least likely of times how remembering the past biases forecasts of the future," *Psychological Science*, 16 (8), 626–630.
- Mynatt, Clifford R., Michael E. Doherty, and Ryan D. Tweney (1977), "Confirmation bias in a simulated research environment: An experimental study of scientific inference," *The quarterly journal of experimental psychology*, 29 (1), 85–95.
- Nedungadi, Prakash (1990), "Recall and consumer consideration sets: Influencing choice without altering brand evaluations," *Journal of consumer research*, 17 (3), 263–276.

- Nickerson, Raymond S. (1998), "Confirmation bias: A ubiquitous phenomenon in many guises.," *Review of general psychology*, 2 (2), 175.
- Nisbett, Richard E. and Timothy D. Wilson (1977), "The halo effect: Evidence for unconscious alteration of judgments.," *Journal of personality and social psychology*, 35 (4), 250.
- Ochsner, Kevin N. (2000), "Are affective events richly recollected or simply familiar? The experience and process of recognizing feelings past.," *Journal of Experimental Psychology: General*, 129 (2), 242.
- Oskamp, Stuart (1965), "Overconfidence in case-study judgments.," *Journal of consulting psychology*, 29 (3), 261.
- Papke, Leslie E. and Jeffrey Wooldridge (1993), *Econometric methods for fractional response variables with an application to 401 (k) plan participation rates*, National Bureau of Economic Research Cambridge, Mass., USA.
- Park, D. C. and Daniel P. Kidder (1996), "Prospective memory and medication adherence," *Prospective memory: Theory and applications*, 369–390.
- Peters, Rik GM and Tammo HA Bijmolt (1997), "Consumer memory for television advertising: A field study of duration, serial position, and competition effects," *Journal of Consumer Research*, 23 (4), 362–372.
- Peterson, Lloyd and Margaret Jean Peterson (1959), "Short-term retention of individual verbal items.," *Journal of experimental psychology*, 58 (3), 193.
- Postman, Leo and Benton J. Underwood (1973), "Critical issues in interference theory," *Memory & Cognition*, 1 (1), 19–40.
- Reason, James T. and Klara Mycielska (1982), *Absent-minded?: The psychology of mental lapses and everyday errors*, Prentice Hall.

- Ross, William T. and Elizabeth H. Creyer (1992), "Making inferences about missing information: The effects of existing information," *Journal of Consumer Research*, 19 (1), 14–25.
- Rottenstreich, Yuval and Amos Tversky (1997), "Unpacking, repacking, and anchoring: advances in support theory.," *Psychological review*, 104 (2), 406.
- Rozenblit, Leonid and Frank Keil (2002), "The misunderstood limits of folk science: An illusion of explanatory depth," *Cognitive science*, 26 (5), 521–562.
- Sanbonmatsu, David M., Frank R. Kardes, and Paul M. Herr (1992), "The role of prior knowledge and missing information in multiattribute evaluation," *Organizational Behavior and Human Decision Processes*, 51 (1), 76–91.
- , ———, David C. Houghton, Edward A. Ho, and Steven S. Posavac (2003), "Overestimating the importance of the given information in multiattribute consumer judgment," *Journal of Consumer Psychology*, 13 (3), 289–300.
- , ———, and Carol Sansone (1991), "Remembering less and inferring more: Effects of time of judgment on inferences about unknown attributes," *Journal of Personality and Social Psychology*, 61 (4), 546.
- Scheibehenne, Benjamin, Rainer Greifeneder, and Peter M. Todd (2010), "Can there ever be too many options? A meta-analytic review of choice overload," *Journal of Consumer Research*, 37 (3), 409–425.
- Schwenk, Charles R. and Richard A. Cosier (1980), "Effects of the expert, devil's advocate, and dialectical inquiry methods on prediction performance," *Organizational Behavior and Human Performance*, 26 (3), 409–424.

- See, Kelly E., Craig R. Fox, and Yuval S. Rottenstreich (2006), "Between ignorance and truth: Partition dependence and learning in judgment under uncertainty.," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32 (6), 1385.
- Shafir, Eldar (1993), "Choosing versus rejecting: Why some options are both better and worse than others," *Memory & cognition*, 21 (4), 546–556.
- , Itamar Simonson, and Amos Tversky (1993), "Reason-based choice," *Cognition*, 49 (1), 11–36.
- Sharpe, William F. (1991), "The arithmetic of active management," *Financial Analysts Journal*, 47 (1), 7–9.
- Shen, Hao and Robert S. Wyer (2008), "Procedural priming and consumer judgments: Effects on the impact of positively and negatively valenced information," *Journal of Consumer Research*, 34 (5), 727–737.
- Simonson, Itamar, Ziv Carmon, and Suzanne O'curry (1994), "Experimental evidence on the negative effect of product features and sales promotions on brand choice," *Marketing Science*, 13 (1), 23–40.
- Soll, Jack B. (1996), "Determinants of overconfidence and miscalibration: The roles of random error and ecological structure," *Organizational Behavior and Human Decision Processes*, 65 (2), 117–137.
- and Joshua Klayman (2004), "Overconfidence in interval estimates.," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30 (2), 299.
- Spiller, Stephen A., Gavan J. Fitzsimons, John G. Lynch Jr, and Gary H. McClelland (2013), "Spotlights, floodlights, and the magic number zero: Simple effects tests in moderated regression," *Journal of Marketing Research*, 50 (2), 277–288.

- Squire, Larry R. (1987), *Memory and brain*, New York: Oxford University Press.
- Tannenbaum, David, Craig R. Fox, and Gülden Ülkümen (2016), “Judgment extremity and accuracy under epistemic vs. aleatory uncertainty,” *Management Science*.
- Teigen, Karl Halvor (1994), “Variants of subjective probabilities: Concepts, norms, and biases.”
- Tulving, Endel and Zena Pearlstone (1966), “Availability versus accessibility of information in memory for words,” *Journal of Verbal Learning and Verbal Behavior*, 5 (4), 381–391.
- Tversky, Amos and Daniel Kahneman (1971), “Belief in the law of small numbers.” *Psychological bulletin*, 76 (2), 105.
- and Derek J. Koehler (1994), “Support theory: A nonextensional representation of subjective probability.” *Psychological review*, 101 (4), 547.
- Ülkümen, Gülden, Craig R. Fox, and Bertram F. Malle (2016), “Two dimensions of subjective uncertainty: Clues from natural language.”
- Vortac, O. U., Mark B. Edwards, and Carol A. Manning (1995), “Functions of external cues in prospective memory,” *Memory*, 3 (2), 201–219.
- Walters, Daniel J., Philip M. Fernbach, Craig R. Fox, and Steven A. Sloman (in press), “Known unknown: A critical determinant of confidence and calibration,” *Management Science*.
- Wang, Jing and Angela Y. Lee (2006), “The role of regulatory focus in preference construction,” *Journal of Marketing research*, 43 (1), 28–38.
- Weber, Elke U., Ann-Renee Blais, and Nancy E. Betz (2002), “A domain-specific risk-attitude scale: Measuring risk perceptions and risk behaviors,” *Journal of behavioral decision making*, 15 (4), 263–290.
- Werth, Lioba and Jens Foerster (2007), “How regulatory focus influences consumer behavior,” *European Journal of Social Psychology*, 37 (1), 33–51.

Wixted, John T. and Ebbe B. Ebbesen (1997), "Genuine power curves in forgetting: A quantitative analysis of individual subject forgetting functions," *Memory & cognition*, 25 (5), 731–739.