http://pubsonline.informs.org/journal/mnsc/

MANAGEMENT SCIENCE
Articles in Advance, pp. 1–10
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

When Risk Is Weird: Unexplained Transaction Features Lower Valuations

Robert Mislavsky,[a] Uri Simonsohn[a]

[a] The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania 19104

Contact: rmisl@wharton.upenn.edu, http://orcid.org/0000-0002-9620-3528 (RM); uws@wharton.upenn.edu, http://orcid.org/0000-0002-8601-7211 (US)

Received: July 15, 2016
Revised: December 30, 2016; May 10, 2017
Accepted: June 5, 2017
Published Online in Articles in Advance: December 4, 2017

https://doi.org/10.1287/mnsc.2017.2868

Abstract. We define transactions as weird when they include unexplained features, that is, features not implicitly, explicitly, or self-evidently justified, and propose that people are averse to weird transactions. In six experiments, we show that risky options used in previous research paradigms often attained uncertainty via adding an unexplained transaction feature (e.g., purchasing a coin flip or lottery), and behavior that appears to reflect risk aversion could instead reflect an aversion to weird transactions. Specifically, willingness to pay drops just as much when adding risk to a transaction as when adding unexplained features. Holding transaction features constant, adding additional risk does not further reduce willingness to pay. We interpret our work as generalizing ambiguity aversion to riskless choice.

History: Accepted by Yuval Rottenstreich, judgment and decision making.
Supplemental Material: The online supplement is available at https://doi.org/10.1287/mnsc.2017.2868 and http://osf.io/fzjuw.

Keywords: transaction features • weirdness • risk aversion • ambiguity aversion • uncertainty effect

The amount people are willing to pay for a given item is influenced by the context in which the purchase takes place (Ariely et al. 2006, Jung et al. 2016, Lichtenstein and Slovic 2006). Transactions, the necessary steps to acquire the item, are a part of every purchase context.

In this paper, we identify a transaction attribute that negatively influences willingness to pay: the extent to which it contains features that lack an explanation. These explanations may be (i) implicit, based directly on consumers' past experiences with similar transactions; (ii) explicit, explained by the seller; or (iii) self-evident, based on reasonable inferences from context. For brevity, we refer to transactions that include unexplained features as "weird." When using the term weird, we refer *exclusively* to such a definition—the presence of unexplained features.

To illustrate how the presence of unexplained features may manifest itself in a transaction and how the three aforementioned types of explanations might mitigate their impact on willingness to pay, consider a restaurant that sells lunches by placing them in boxes and then asks people to pay for the right to open the box and take the lunch from the box. Placing the lunch in a box and asking to pay to open it could constitute an unexplained transaction feature and may make customers uncomfortable or suspicious (e.g., is the lunch in the box because the restaurant doesn't want you to see what it really looks like?). An explanation could easily mitigate any such consequences. An *implicit* explanation would be if the "box" was simply a vending machine; customers could draw on their prior experience and the transaction feature is no longer unexplained. Alternatively, the restaurant could provide an *explicit* explanation "These are our new self-service boxes, which we've introduced to help you get your food more easily." The transaction feature is again no longer unexplained and thus no longer expected to reduce valuations.

Note that unexplained is not the same as novel. A completely novel transaction feature could come with an explanation. For example, imagine a restaurant that requires customers to draw on a piece of glass with their finger to get their lunch. That is an unusual transaction feature. But if the glass is an iPad screen, and the drawing is the customer's signature, customers facing this transaction feature for the very first time would easily generate a *self-evident* explanation for why the transaction feature is there. It would not be expected to lower valuations.

We conjecture that the presence of unexplained features lowers willingness to pay because it triggers reactions akin to ambiguity aversion (Ellsberg 1961, Frisch and Baron 1988, Keren and Gerritsen 1999) in general and comparative ignorance in particular (Chow and Sarin 2001, Fox and Tversky 1995, Fox and Weber 2002). Relevant but unknown information may make consumers less confident in the decision to make the purchase (Chow and Sarin 2001, Fox and Tversky 1995, Fox and Weber 2002) or perhaps make them feel the seller has more information that she may use to her advantage (Frisch and Baron 1988, p. 153; Keren and Gerritsen 1999). The presence of unexplained features

creates an imbalance between seller and buyer in terms of what relevant information they have for the transaction. Weirdness aversion, the aversion to transactions with unexplained features, may then constitute the generalization of ambiguity aversion to situations that lack (explicit) uncertainty.

We demonstrate the practical relevance of an aversion to unexplained transaction features by focusing on a research paradigm where researchers unintentionally manipulated the presence of unexplained transaction features and obtained a result, often referred to as the "uncertainty effect" (Gneezy et al. 2006). We find that the uncertainty effect may instead be caused by weirdness, or the presence of unexplained transaction features.

Gneezy et al. (2006) documented that people were willing to pay less for a risky prospect than for its worst possible outcome. For instance, people were willing to pay an average of \$26.10 for a \$50 Barnes and Noble gift card but only \$16.12 for a gamble where participants were guaranteed to win either a \$50 or \$100 gift card, each with a 50% probability. This general finding has been replicated by many independent research teams (e.g., Andreoni and Sprenger 2011, Newman and Mochon 2012, Simonsohn 2009, Wang et al. 2013, Yang et al. 2013).[1]

These uncertainty effect studies pit valuations of a risky option against valuations of a riskless one. The risky option requires a mechanism that introduces risk, while the riskless option does not. For example, researchers have generated risky prospects by asking participants to buy coin flips, lottery tickets, unlabeled envelopes, and gift cards of unknown value and have compared participants' valuations of these transactions to that of buying a gift card outright. There is no explicit or implicit justification to sell gift cards of unknown value or to utilize a coin flip to determine their value. Therefore, while these mechanisms do generate risk, they also introduce unexplained features to the transaction.

Uncertainty effect studies, therefore, have included a risky transaction *with* unexplained features and a not risky one *without* unexplained features, perfectly confounding risk with weirdness. In this paper, we report studies that manipulate the presence of unexplained features independently of risk. Our results are consistent with an aversion to unexplained features accounting for somewhere between the preponderance and the totality of the uncertainty effect. After presenting our empirical results, we discuss how unexplained features could be present in other paradigms used to study consumer behavior.

Transparent Reporting

Studies 1–5 were run on Amazon's Mechanical Turk (MTurk) and were administered through Qualtrics. Study 6 was incentive compatible and run in a behavioral lab. For all studies we decided sample size before collecting any data. MTurk participants were not allowed to participate in more than one study. We included attention checks for Studies 5A and 5B. Studies 5A, 5B, and 6 were preregistered. For all studies we report all data exclusions (if any), all manipulations, and all measures. Data, analysis code, preregistrations, and survey materials are available at http://osf.io/x8cqm.

Studies 1–3: Weird, But Not Risky

Our first three studies are similar, so we present them together. In all three we modified the traditional uncertainty effect paradigm to disentangle the effect of risk from the effect of unexplained features on valuations. For a more fluent reading experience, we refer to transactions that include unexplained features as "weird" and to the presence of such features as "weirdness." The uncertainty effect paradigm pits the valuation of a riskless prospect (e.g., buying a \$50 Target gift card) against that of a risky one (e.g., flipping a coin to determine if the gift card is for Target or for Walmart). This paradigm confounds risk and weirdness because the manipulation that introduces risk also introduces unexplained features to the transaction (e.g., flipping a coin). To examine the importance of this confound, we created a third type of transaction, one that was *weird but not risky*. Specifically, this was a transaction that includes the same unexplained features present in the risky transactions (e.g., buying a token redeemable for a gift card) but with a certain outcome (e.g., the value of the gift card is known).

Method

Design. In Study 1 (N = 603; 29.6% female), we randomly assigned participants to one of three conditions asking them indicate their maximum willingness to pay (WTP) for a transaction. The first two were analogous to traditional uncertainty effect studies:

Condition 1. *Neither weird nor risky*[2]

"We want to know how much you would be willing to pay for two different items, a \$50 Walmart gift card and a \$50 Target gift card.
If you could buy only the \$50 Walmart gift card, what is the most you would pay for it?___
If you could buy only the \$50 Target gift card, what is the most you would pay for it?___"

Condition 2. *Weird and risky*

"Imagine that you are standing in front of a table that has a locked box on it. The box has two gift cards inside: a \$50 Walmart and a \$50 Target gift card.
You can pay to open the box and choose a gift

card, which will be yours to keep. The gift cards do not have the names of the stores printed on them, so you will not know which gift card is which.
What is the most you would be willing to pay to open the box?___"

Uncertainty effect studies compare the valuation of similar pairs of transactions. Any difference in WTP can therefore be caused by the risk difference (having a known versus unknown outcome) or by the weirdness difference (buying outright versus paying to open a box). We addressed this confound by adding a *weird but not risky* condition. Participants read the same scenario as those in the *weird and risky* condition, except the gift cards were labeled, so participants knew which card they were getting before choosing. Specifically, it read (differences between Conditions 2 and 3 underlined here but not in original materials):

Condition 3. *Weird but not risky*
"Imagine that you are standing in front of a table that has a locked box on it. The box has two gift cards inside: a $50 Walmart and a $50 Target gift card.
You can pay to open the box and choose a gift card, which will be yours to keep. The gift cards have the names of the stores printed on them, so you will know which gift card is which.
What is the most you would be willing to pay to open the box?___"

After running this study, we identified a potential confound. The weird transactions (paying to take one of two gift cards from a box) had two possible outcomes, while the not weird transaction had only one. We believed this difference, rather than weirdness, could explain any observed differences (e.g., because people are averse to explicitly rejecting an outcome). In Study 2 ($N = 308$; 35.5% female) we reran the two weird conditions and added a new weird condition that had only one possible outcome. Across the three conditions, then, participants paid to open a box and take a card from it. The conditions differed on whether the box contained *one labeled* gift card (new condition), *two labeled* gift cards, or *two unlabeled* gift cards. We did not rerun the *neither risky nor weird* condition.

In Study 3 ($N = 403$; 36.8% female) we reran all four conditions from Studies 1 and 2 with a different operationalization of risk and weirdness: purchasing a token at an event and redeeming it for a gift card. The four conditions were as follows:

1. *Neither weird nor risky*
"What is the highest amount you would be willing to pay for a $50 [Walmart/Target] gift card?"
(*Target and Walmart counterbalanced within-subjects*)

2. *Weird but not risky, one option*
"Imagine that you are at an event where there are tokens for sale. These tokens can be redeemed at a cashier for a $50 [Walmart/Target] gift card. What is the highest amount you would be willing to pay for one of these tokens?"
(*Target and Walmart counterbalanced within-subjects*)

3. *Weird but not risky, two options*
"Imagine that you are at an event where there are tokens for sale. These tokens can be redeemed at a cashier for your choice of either a $50 Walmart gift card or a $50 Target gift card. What is the highest amount you would be willing to pay for one of these tokens?"

4. *Weird and risky*
"Imagine that you are at an event where there are tokens for sale. These tokens can be redeemed at a cashier for either a $50 Walmart gift card or a $50 Target gift card. The cashier will flip the token, and if it lands on heads, you will receive the Walmart gift card. If it lands on tails, you will receive the Target gift card. What is the highest amount you would be willing to pay for one of these tokens?"

Results

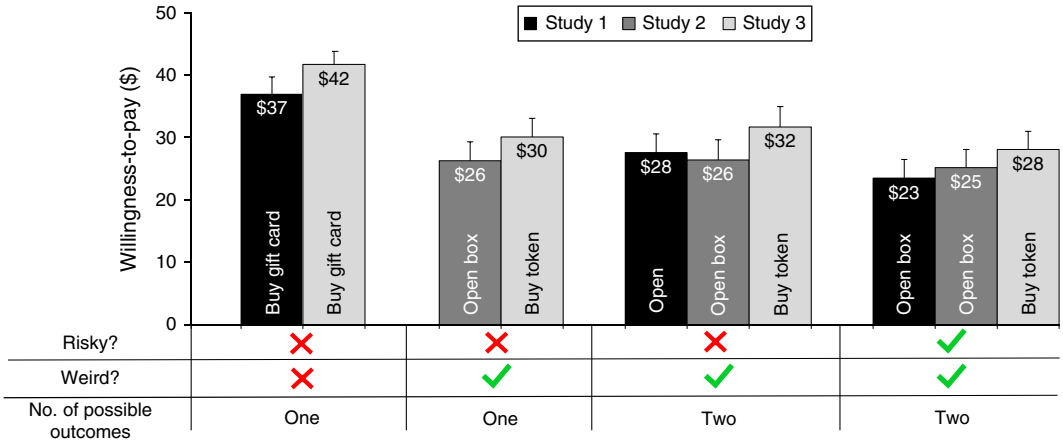
Figure 1 depicts results for Studies 1–3. We identify four main takeaways:

1. In Studies 1 and 3, we replicate the original uncertainty effect (Study 2 does not allow testing it). Participants valued the *weird and risky* prospects ($M = \$25.80$), less than their least-valued *neither weird nor risky* gift card ($M = \$39.37$). The risky option was valued significantly less than its worst outcome in both studies, ts > 6.58, ps < 0.001.

2. Holding weirdness constant, there is no apparent uncertainty effect. Comparing the two weird conditions, *risky* gift cards ($M = \$25.60$ across all studies) were not valued significantly less than the *riskless* gift cards ($M = \$28.39$ across all studies), whether they had one or two options (Study 1: $t(199) = 1.92$, $p = 0.057$; Study 3: ts < 1.64, ps > 0.10). Based on point-estimates of the means, the effect of weirdness is two-thirds (Study 1) to three-quarters (Study 3) as large as the uncertainty effect is when weirdness is not accounted for.[3] We believe some of this residual effect we are attributing to uncertainty is also attributable to weirdness, because it seems likely that, in these scenarios, uncertainty makes the weird scenarios weirder by adding an additional unexplained feature (e.g., flipping a token to determine the value of a gift card). We could not estimate this for Study 2, because it did not include a *not weird* condition.

3. Contrary to our initial expectations, these results are not driven by the number of potential options.

Figure 1. (Color online) Average Valuations (Studies 1–3) as a Function of Risk and Weirdness



Notes. Hypothetical valuations for $50 gift cards. Risk involves whether it is for Target or Walmart, operationalized via opening a box and selecting one of two *unlabeled* envelopes (Studies 1 and 2), or purchasing a token exchangeable for one of the two gift cards, determined by flipping the token (Study 3). Weird but riskless involves labeled envelopes (Studies 1 and 2), or participants *choosing* what to redeem the token for (Study 3). Transactions with one outcome (bottom row) involve box with one gift card (Study 2) or token with predetermined value (Study 3). Error bars represent 95% confidence intervals.

Valuations for the weird but not risky transactions are similar when they involve one or two possible outcomes, ts < 0.71, ps > 0.47.

4. Study 3 rules out a potential confound for Studies 1 and 2. In the box scenarios, participants may have believed that they had to make two payments, one to open the box and another to purchase the gift card. Because very few participants paid $0 in the weird scenarios (as would be expected if this were the case; see Online Supplement 2), we believe this is unlikely, although a reviewer also raised the possibility that participants may have averaged the two payments when reporting their WTP. We obtain very similar results in the token scenario, where this ambiguity is not present, which appears to rule this possibility out.

Study 4: Bigger Differences in Outcomes

In the first three experiments, the risky prospects involved gift cards with the same face value ($50) for different stores (e.g., Target versus Walmart). This design, originally used by Newman and Mochon (2012), allowed us to create *weird but not risky* conditions where participants could meaningfully choose between gift cards, whereas choosing between a $50 card and a $100 card is not a meaningful choice. However, minimizing outcome variance may have inflated the importance of the unexplained features. In other words, we may have found risk did not matter much because we created situations without much risk. In this experiment, we created risky prospects with greater outcome variance.

Method

Sample. We recruited 604 participants (39.4% female), each paid $0.25.

Design. Participants were randomly assigned to one of eight conditions in a between-subjects design. Two *not weird* conditions were similar to those in Studies 1 and 3: participants provided their WTP for either a $50 Target gift card or a $100 Target gift card bought outright. The remaining six conditions involved *weird* transactions and conformed to a 2 (*transaction*: box versus token) × 3 (*value*: $50 versus $100 versus risky) design. Participants read either the box or token scenarios from the prior studies, where the outcomes were either a $50 Target gift card for sure, a $100 Target gift card for sure, or a Target gift card that was worth either $50 or $100, each with 50% probability. We did not include a condition where participants could choose either a $50 or $100 gift card because we assumed all participants would choose $100. We decided before data collection began to obtain 120 observations from the *not weird* conditions and 60 from each *weird* condition (since we had two versions of weirdness, $60 * 2 = 120$).

Results

Beginning with the token conditions, the uncertainty effect was again replicated when not accounting for transaction weirdness. Participants valued the risky token $6.27 less than they did its worst possible outcome purchased outright ($M = \$37.23$ and $M = \$43.50$, respectively), $t(179) = 2.70$, $p = 0.008$. Comparing the weird conditions, people paid $5.59 *more* for the risky prospect (*token exchangeable for* $50 *gift card*: $M = \$31.64$; *risky token*: $M = \$37.23$), $t(118) = 1.57$, $p = 0.12$.[4]

The uncertainty effect was also replicated in the box conditions ($50 *Target gift card bought outright*: $M = \$43.50$; *risky box*: $M = \$25.23$), $t(180) = 9.13$, $p < 0.001$. The difference between the risky prospect and its least valued outcome was much smaller when comparing

the two weird conditions (\$50 *gift card in box*: M = \$29.44; *risky box with* \$50 *or* \$100 *gift card*: M = \$25.23), $t(120) = 1.40$, $p = 0.16$. The total uncertainty effect is about \$18 (\$43.50 − \$25.23). The effect of weirdness alone is about \$14. As argued above, the residual \$4 effect could be the result of weirdness if choosing among unlabeled cards seems less justified than taking a labeled card out of a box.

There was also a sizable main effect of weirdness for individual valuations of the \$50 and \$100 gift cards. Buying a \$50 or \$100 gift card outright was valued at \$43.50 and \$86.49, respectively, whereas a \$50 or \$100 gift card in a box was valued at \$29.44 and \$51.47, respectively, and a token exchangeable for a \$50 or \$100 gift card was valued at \$31.64 and \$65.93, respectively, ts > 5.38, ps< 0.001. We report all pairwise comparisons in Online Supplement 4. In sum, we obtain results similar to those of Studies 1–3 using risky prospects with greater outcome variance. The data are consistent with unexplained features accounting for somewhere between the preponderance and the totality of the uncertainty effect.

Studies 5A and 5B: Evaluating Weirdness of Prior Uncertainty Effect Studies

In Study 5 we more directly test if prior uncertainty effect studies have unintentionally manipulated weirdness by asking participants to evaluate the weirdness of the underlying transactions in those studies.

One may measure weirdness on absolute or relative scales, although each has its limitations. Absolute scales (e.g., "How weird is this transaction?") are ambiguous about what a transaction is being compared to, or equivalently, what the values in the scale represent. Relative scales, on the other hand, (e.g., "Which transaction is weirder?"), may create demand effects or change participants' definitions of weirdness where they think the weirdest transaction is the one that is least like the others (even though it may be the simplest). Since neither approach was obviously superior, we pursued both, and in both cases we explicitly defined weirdness to our participants as involving the presence of unexplained features. Participants judged weirdness on both an absolute scale (Study 5A) and on a relative scale (Study 5B). We obtained consistent results with both methods. Risky transactions in prior uncertainty effect studies are weirder than their riskless counterparts.

Study 5A: Between-Subjects Ratings of Weirdness

Method

Sample. We recruited 714 MTurk participants, 600 of whom (53.3% female, M_{age} = 35.3 years) passed an attention check and were able to continue to the rest of the survey, each paid \$0.40 (preregistration: https://aspredicted.org/3mu9d.pdf).

Design. In a between-subjects design, participants evaluated the weirdness of transactions used in prior uncertainty effect studies. Participants began by reading this passage:

> "We will show you an example of a purchase that experimenters ask participants to evaluate. We are interested in knowing how 'weird' you think the purchase is. By 'weird,' we mean how much the purchase has unusual and unexplained features."

Participants then read one of eight questions used in prior uncertainty effect studies—two from Gneezy et al. (2006), three from Yang et al. (2013), and three from this paper. Three of these questions were "baseline" questions (i.e., the riskless valuations that were used as control conditions in uncertainty effect studies).[5] We preregistered that we would collapse the ratings for these conditions for analysis. The other five valuations were used in prior studies—Gneezy et al. (2006, p. 1301) lottery, Yang et al. (2013, p. 737) certain and uncertain coins, and our certain and uncertain boxes (Study 4). See the appendix for the exact text of these stimuli. After reading the question, participants rated its weirdness using the following scale: "How weird is it to buy a gift [card/certificate] like this?" (1 = It is not weird at all; 2 = It is a little weird; 3 = It is very weird; 4 = It is extremely weird). If risk and weirdness were confounded in these studies, we would expect that the weird transactions would be rated as weirder than the baseline ones.

Results

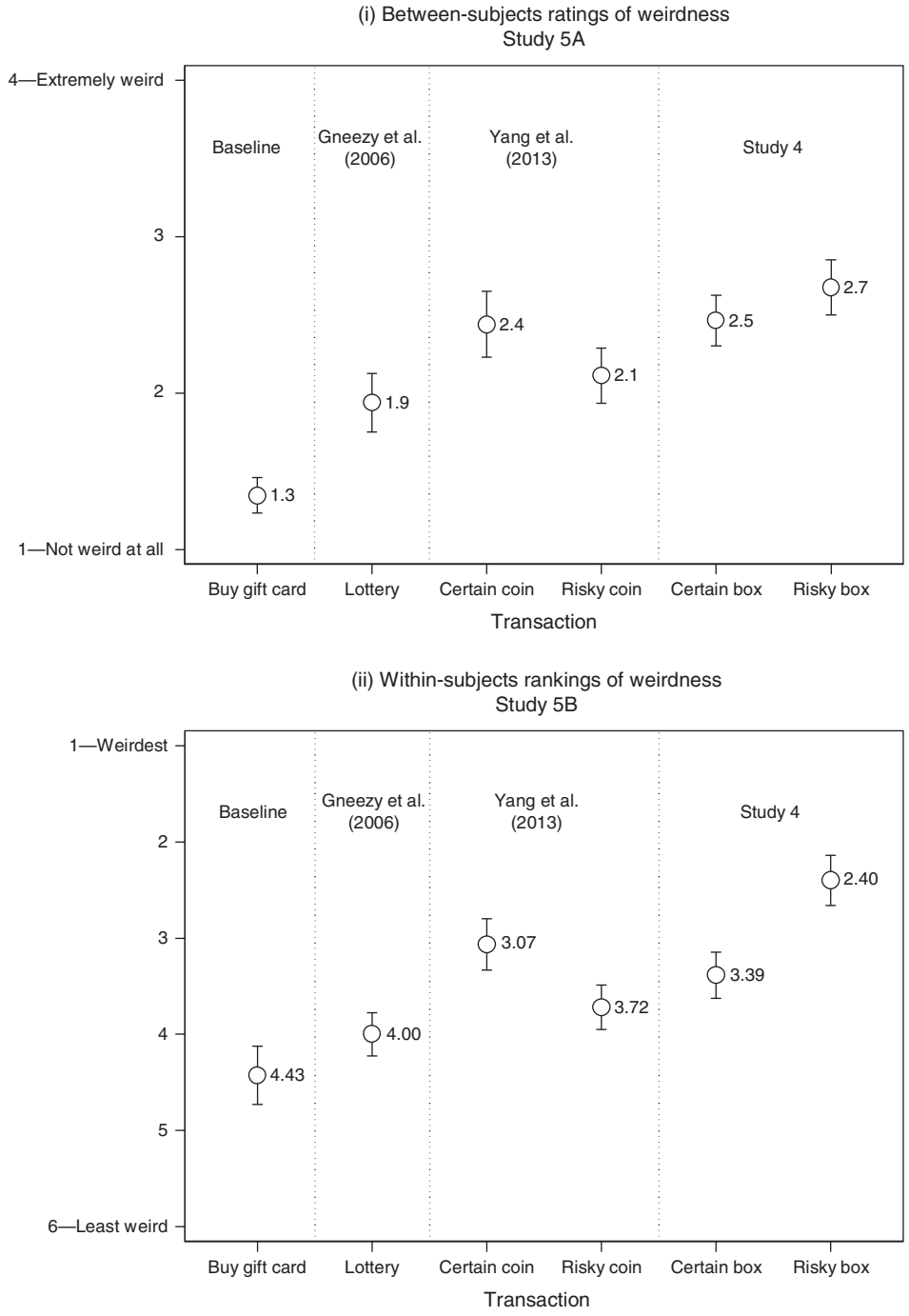
Consistent with the notion that prior uncertainty effect studies have confounded risk and weirdness, participants rated all of the weird transactions ($1.94 \leq M$s ≤ 2.68) as weirder than the baseline transaction (M = 1.35), all ts > 4.83, all ps < 0.001. See Figure 2, panel (i). In addition to this preregistered comparison, we compared the share of participants rating a transaction as "not weird at all." Seventy percent of participants gave this rating to the baseline transaction compared to between 8% and 39% for the weird transactions, Zs > 4.45, ps < 0.001.

Study 5B: Within-Subjects Rankings of Weirdness

Method

Sample. We recruited 184 participants, 153 of whom (42.7% female, M_{age} = 35.5 years) passed an attention check and were able to continue to the rest of the survey, each paid \$0.40 (preregistration: https://as-predicted.org/p4hi5.pdf).

Figure 2. Prior Uncertainty Effect Studies Are Weirder Than Their Baseline Comparisons



Notes. Panel (i) shows between-subjects ratings (Study 5A; $N = 600$) of transactions used in prior uncertainty effect studies (see the appendix). The scenarios were described verbatim to participants. The y axis shows the average response to the question: "*We are interested in knowing how 'weird' you think the purchase is . . . By 'weird,' we mean how much the purchase has unusual and unexplained features.*" Panel (ii) shows within-subjects rankings of weirdness (Study 5B; $N = 153$) of the same scenarios.

Design. All participants were given the same instructions as in Study 5A, but instead of rating them between-subjects, they were shown six transactions (one of the three baseline transactions and all five weird transactions) and asked to *rank* them from weirdest (1) to least weird (6). Ties were not allowed.

Results

Consistent with Study 5A and more generally with the notion that prior uncertainty effect studies have confounded risk and weirdness, participants ranked purchasing a gift card outright as the least weird ($M = 4.43$ out of 6) out of all the transactions (between 2.40 for the

Risky Box, $t(149) = 8.47$, $p < 0.001$, and 4.00 for the GLW Lottery, $t(149) = 2.00$, $p = 0.047$). See Figure 2, panel (ii). Here the weirdness difference between the baseline and the original uncertainty effect (Gneezy et al. 2006) seems smaller than in Study 5A. Part of this may be explained by some participants reversing the scale, since 14% of participants ranked the baseline transaction as the *weirdest* (the second most popular answer). Nevertheless, looking at the number of participants who ranked the transaction as least weird, a comparison not included in our preregistration, we see a more substantial difference. Specifically, while 46% of people ranked the baseline as the least weird, only 15% did for the Gneezy et al. (2006) lottery, $Z = 5.97$, $p < 0.001$.

Study 6: Incentivized Lab Study

To this point, all of our studies have used hypothetical scenarios. To address the possibility that our findings were driven in part by participants' inattention or lack of motivation, our last study is an incentive-compatible replication (preregistration: https://aspredicted.org/dq97y.pdf).

Method

Sample. We recruited 219 participants (71.1% female, $M_{age} = 20.8$ years) at the Wharton Behavioral Lab. This study was part of a larger lab session with several unrelated studies, and all participants were paid $10 for completing the session

Design. In a three-cell between-subjects design, participants indicated their willingness to pay (WTP) for an item. The three conditions were (i) buying a $50 Amazon gift card (*neither weird nor risky* condition), (ii) paying to open a locked box with a $50 Amazon gift card and taking the card (*weird but not risky* condition), and (iii) paying to open a locked box containing a $50 gift card and a $100 gift card, with values only visible on the inside, and taking a card without knowing its value (*weird and risky* condition).

One in every 20 participants was randomly selected to have their decision count for real and receive a $100 bonus (to fund the purchase). To make the WTP elicitation incentive-compatible, a price was set but not revealed to participants. If participants' WTP was greater than that price, they made the purchase and paid that price. Otherwise, they kept the entire bonus and did not make a purchase. To indicate their WTP, we showed participants a price, starting at $5, and they indicated if they would make the purchase for that amount. If they said yes, we increased the price by $5, and they answered again. This was repeated until they answered "No" or the price reached $100.[6] The highest price participants said "Yes" to is our dependent variable. We purposefully avoided a multiple price list and used a multiple price *sequence*, concerned that the price list could prompt participants to choose valuations in the middle of the range for the uncertainty condition, attenuating the uncertainty effect (original materials: https://osf.io/kty32).[7]

Results

Without accounting for weirdness, for the presence of unexplained transaction features, participants again acted as if they were extremely risk averse. Willingness to pay for the *weird and risky* transaction ($M = \$39.24$) was similar to that for the *neither weird nor risky* one ($M = \$38.70$), $t(143) = 0.19$, $p = 0.85$, even though the former has an expected value approximately 50% higher than the latter. As in prior uncertainty effect studies, this suggests the presence of direct risk aversion, since neither prospect theory nor expected utility theory can generate such extreme levels of risk aversion. But if defined narrowly, as obtaining a strictly lower mean, this result does not replicate the uncertainty effect.[8] In any case, this comparison confounds risk and weirdness.

Controlling for weirdness, participants appear to show very mild (if any) risk aversion: the risky purchase ($M = \$39.24$) was valued noticeably *above* the not risky one ($M = \$30.47$), $t(143) = 2.94$, $p = 0.004$. In fact, participants valued the uncertain gift card close to what a risk *neutral* buyer would be expected to value it. In particular, assuming participants would pay twice as much for a $100 gift card as they would for a $50 gift card (which is a conservative assumption that does not account for diminishing sensitivity or marginal utility), a risk neutral valuation of the risky gift card is $45.71 ($1.5 * \$30.74$), which is not much higher than what we observe ($39.24), $t(143) = 1.77$, $p = 0.080$.[9]

Finally, holding risk constant, we replicate weirdness aversion. The not weird purchase ($M = \$38.70$) was valued above the weird one ($M = \$30.48$), $t(144) = 3.17$, $p = 0.002$.

General Discussion

We have documented that the presence of unexplained features lowers willingness to pay (WTP). We manipulated the presence of such features, weirdness, independently of risk and found that the effect of weirdness on WTP is of about the same magnitude as the uncertainty effect, which had previously been attributed to the presence of uncertainty. These results suggest that subtle transaction features can have dramatic effects on WTP—dramatic enough for multiple independent research teams to run successful replications of the original Gneezy et al. (2006) finding but subtle enough that they did not notice the potential confound when doing so (including one of us; see Simonsohn 2009).

Unexplained Features Is the Key Manipulation

We have characterized our key manipulations as increasing weirdness, or introducing unexplained features to transactions. Some of the seven members of our review team proposed alternative interpretations for our manipulations. One reviewer proposed that

perhaps we simply manipulated the total number of features (whether weird or not). We do not believe the number of features per se is critical. First, in an experiment included in a prior version of the manuscript, we found that merely adding features did not reduce valuations (see Online Supplement 6). Second, in many empirical studies, valuations are often elicited with procedure that require different numbers of steps (e.g., asking for a price outright versus going through a multiple price list), and it has not been previously documented that transactions with more steps lead to lower valuations. Third, there is no obvious psychological process that would seem to justify this prediction. In contrast, we believe that all mechanisms that have been proposed for ambiguity aversion would also predict that *unexplained* features lower valuations.

Another reviewer proposed that perhaps what is special about the features we introduced is not that they are unexplained features, but that they are unusual features that transactions outside the lab would not include. That is to say, people would pay less for opening a box to buy an item, not because they see no reason to have that extra step, but because outside the lab they have never purchased an item by paying to open the box. We do not find this alternative explanation compelling either. First, most transactions in the lab are rather unusual. Take, for example, our baseline condition in incentive-compatible Study 6. Participants completed a multiple price sequence, which was then compared with a preset price to determine if they would purchase a $50 gift card held by the experimenter. This is not a transaction they would engage in outside the lab. And yet, their WTP was a rather high $38.70 and comparable to the valuations from prior studies that did not involve the convoluted incentive-compatible mechanism (e.g., $37 in Study 1 here).

Second, we can easily imagine situations where a completely new transaction feature, because it is accompanied by an explanation, would not be expected to lower WTP. Consider again that example from the introduction about a person's first payment by signing on an iPad, or perhaps an American asked to pay in rubles during her first coffeeshop visit in Moscow. In these examples, consumers are facing entirely novel transaction features, but these features have self-evident explanations and would not be predicted to lower WTP.

When Risk Is Not Weird

Our studies manipulate unexplained features independently of risk (i.e., we include transactions that are *weird but not risky*), but not risk independently of unexplained features (i.e., we do not include transactions that are *not weird but risky*). The absence of a *not weird but risky* cell in our experiments may pose some problems for the interpretation of our studies. If a *not weird but risky* condition was valued similarly (or lower) than a *weird and risky* scenario, it would imply that unexplained features moderate, rather than account for, the effect of risk in those transactions. Although we think this is unlikely, our data cannot rule this out.

This is a challenge to explore empirically because it requires a situation where risk is an expected feature (e.g., buying stocks), and is therefore not weird. In such situations, however, offering an option with no risk (e.g., a riskless stock) would be weird, since it would involve the presence of a feature that requires an explanation ("why is this stock riskless?"). Experiment 4 of Yang et al. (2013) provides an example of our concern. They include a condition where participants indicate their WTP for a coin flip that paid a $50 gift certificate if the coin landed on heads *or* tails ("Certain Coin Flip," p. 737). In our Studies 5A and 5B, we asked participants to rate how weird this transaction was, and they rated it as *weirder* than the risky coin flip (i.e., as containing more unexplained features), likely because a coin flip implies risk and removing risk makes the coin flip unnecessary.

Further, even holding all features of a transaction constant, all risk per se may not be equally unexplained. For instance, in most gambling situations, payoffs are inversely proportional to the probability of winning. Therefore, a lottery with a 1% chance of winning $100 and a 99% chance of winning $50 is more typical (i.e., has an implicit explanation) than a gamble with a 99% chance of winning $100 and a 1% chance of winning $50. If this were true, and if unexplained features reduce valuations, people should appear more risk averse for the latter lottery. A closer look at Gneezy et al. (2006, p. 1287) reveals evidence consistent with this conjecture. Participants are risk seeking (i.e., WTP > Expected Value) when there is a 1% chance of winning the larger price and risk averse (i.e., WTP < Expected Value) when there is a 99% chance of winning the larger prize (p. 1287, table 1). In fact, the median WTP for these two gambles are identical ($37.50) in this study. Of course this is speculative, and there are several potential explanations for these findings that have little to do with the specific transaction features (e.g., probability weighting; McGraw et al. 2010, Rottenstreich and Hsee 2001).

Attributing the Uncertainty Effect to Unexplained Transaction Features May Reconcile Inconsistent Findings

The "direct risk aversion" explanation for the uncertainty effect (Gneezy et al. 2006, Simonsohn 2009) seems at odds with studies that show consumers responding more favorably to risky promotions than to riskless ones. Specifically, Mazar et al. (2017) find that consumers prefer a probabilistic discount to a certain discount of the same expected value (e.g., a 10% chance of getting item for free versus a certain 10% discount),

Table 1. Index of Supplementary Materials

Section	Pages
Supplement 1. Complete age data for Studies 1–4	2
Supplement 2. Additional analyses for Study 1	3–4
Supplement 3. Within-subject variation in valuation of gift cards in Studies 1–3	5
Supplement 4. Pairwise comparisons across all conditions in Study 4	6
Supplement 5. All means and pairwise comparisons for Studies 5A and 5B	7
Supplement 6. Study S1—Isolating and mediating with weirdness	8–10
Supplement 7. Study S2—Comparing ascending and descending price sequences	11

Note. The online supplement is available at https://doi.org/10.1287/mnsc.2017.2868 and http://osf.io/fzjuw.

while Goldsmith and Amir (2010) find that offering a randomly determined prize for making a purchase is nearly as effective as offering the most attractive prize for sure.

If the uncertainty effect were caused by unexplained transaction features, rather than direct risk aversion, at least two explanations arise for the apparent contradiction. First, it may be that consumers can readily identify a reason for a company to offer the type of promotions examined in those studies. They have an explanation, so they are not aversive.[10] Second, in uncertainty effect studies, the focal item (e.g., the gift card participants are purchasing) is uncertain, while in the risky promotion studies, the "bonus" is uncertain. The focal transaction does not contain an unexplained feature; the bonus does. Perhaps people tolerate (or even prefer) these features in such circumstances.

Another difference is that uncertainty effect studies typically use WTP as their dependent variable, while the risky promotion studies use choice (Mazar et al. 2017) and attractiveness ratings (Goldsmith and Amir 2010). Perhaps the WTP question implicitly forces a transaction on participants, enhancing the negative suspicions of buyers, but this pressure dissipates in the other tasks. Moon and Nelson (2017) do not replicate the uncertainty effect with a choice task, but Gneezy et al. (2006, p. 1292) do. The role of elicitation mode on the effects of risk and of unexplained features remains an open question, as there are too many differences in these respective designs to meaningfully interpret the differences in results.

Potential Transaction Feature Confounds in Other Literatures

Much of consumer research involves the comparison of valuations of the same item across different transaction contexts. For example, the endowment effect compares valuations of items being sold against those being purchased, and time preference studies compare the valuations of delayed payments occurring at different points in time (e.g., payments happening today versus payments happening in the future). Those contextual differences may unintentionally have added unexplained features as well.

For example, it may be the case that giving participants an item and immediately ask them to sell it is an atypical feature, relative to giving them money and offering the opportunity to buy an item. This would depress WTP relative to WTA. Similarly, delaying a payment due today may be perceived as less justified than delaying a payment occurring in the future. This potential confound would lead to more severe discounting of immediate than future delays, typically interpreted as evidence of impatience. In many cases, however, controlling for these differences may be difficult. In our case, for example, we could not find a way to induce risk without adding transaction features, so we added features to the riskless option; this may be the easiest path to control for the weirdness confound in other paradigms as well.

This paper contains an online supplement. Table 1 summarizes its contents.

Acknowledgments

The authors thank the review team for constructive comments. The authors also thank the Wharton Risk Management and Decision Processes Center and the Wharton Behavioral Lab for financial support.

Appendix. Stimuli Used in Studies 5A and 5B

Baseline (Randomly Selected from the Following)

- Imagine that you could buy a $50 gift certificate to Barnes and Noble as part of this study. The gift certificate is good for use within the next two weeks.
- Imagine that you could buy a $50 Target gift card as part of this study.
- We are interested in how much you would pay for a $50 Barnes and Noble gift certificate, which you could buy as part of this study.

Gneezy et al. (2006, p. 1301) Lottery

"Imagine that we offer you a lottery ticket that gives you a 50 percent chance at a $50 gift certificate for Barnes and Noble, and a 50 percent chance at a $100 gift certificate for Barnes and Noble. Whichever gift certificate you win is good for use within the next two weeks."

Yang et al. (2013, p. 737) Certain Coin

"We are interested in how much you would be willing to pay for participating in a coin flip. If heads comes up, you will get a $50 gift certificate for Barnes and Noble bookstore. If tails comes up, you will get a $50 gift certificate for Barnes and Noble bookstore."

Yang et al. (2013, p. 737) Uncertain Coin

"We are interested in how much you would be willing to pay for participating in a coin flip. If heads comes up, you will get a $50 gift certificate for Barnes and Noble bookstore. If tails comes up, you will get a $100 gift certificate for Barnes and Noble bookstore."

Study 4 Certain Box

"Imagine that you are standing in front of a table that has a locked box on it. The box has a $50 Target gift card inside. You can pay to open the box and take the gift card, which would be yours to keep."

Study 4 Risky Box

"Imagine that you are standing in front of a table that has a locked box on it. The box has two gift cards inside: a $50 Walmart and a $50 Target gift card.

"You can pay to open the box and choose a gift card, which will be yours to keep. The gift cards do not have the names of the stores printed on them, so you will not know which gift card is which."

Endnotes

[1] Keren and Willemsen (2009) report results where the uncertainty effect is not observed when comparing average valuations. Gideon Keren shared the raw data from that article with us. We analyzed it as in Simonsohn (2009), comparing the entire distributions of responses, and found that a substantial share of participants do show the effect. Rydval et al. (2009) provide the only failure to replicate the uncertainty effect that we are aware of. Their favored explanation is that participants in other experiments misunderstood the task and/or payoffs. Yang et al. (2013) find that the uncertainty effect is only observed for willingness to pay and not for willingness to accept measures.

[2] In Study 1, some participants valued Walmart/Target gift cards and others valued Amazon/Barnes and Noble gift cards. Because subsequent studies only included the former, we report results for the latter in Endnote 3. We also collected data on self-reported average expenditures in other purchases to use as covariates to increase power, but they were uncorrelated with the dependent variable and therefore not useful. We did not collect these in subsequent studies. See Online Supplement 2 for covariate results.

[3] For the Barnes and Noble and Amazon gift cards in Study 1, the means are $35.91 (*neither weird nor risky*), $27.77 (*weird but not risky*), and $22.30 (*weird and risky*). The total uncertainty effect amounts to $13.61, with weirdness accounting for nearly 60% of the effect.

[4] Analyzing the data as in Simonsohn (2009), the lower bound of people paying less for the uncertain item is 3.3% in the *token* conditions and 19.7% in the *box* conditions, neither of which is significantly greater than 0 ($ps > 0.09$). See Online Supplement 4.

[5] These questions were slightly adapted to sound like an actual transaction (e.g., "Imagine you are buying this") rather than an abstract valuation (e.g., "What is the most you are willing to pay for this?").

[6] Only one participant (in the *neither weird nor risky* condition) gave a WTP of $100.

[7] A reviewer expressed this concern about a multiple price sequence that we thought was worth sharing with readers: "[A] price-sequence may not be innocuous, either: The initial, low prices may serve as anchors for subjects' valuations [. . .] which may bias WTPs down. If such anchoring effects were asymmetric, and were more pronounced for risky or weird transactions (because, say, preferences for risky or weird transactions are less stable), then they could make the experimental results difficult to interpret." To respond to this concern, we ran a study on MTurk manipulating whether the multiple price sequence was increasing or decreasing. The effect of weirdness is significant and of the same magnitude for both. See Online Supplement 7.

[8] Although we preregistered that we would calculate the proportion of the uncertainty effect explained by weirdness, we could not do this here because we do not directionally replicate the original uncertainty effect.

[9] To perform this t-test, we multiplied all valuations in the weird but not risky condition by 1.5 and conducted a standard difference of means t-test comparing this new variable with the observed valuations in the *weird and risky* condition. The comparison, therefore, treats $45.71 as an estimated magnitude with a standard error (which it is), rather than as a preset constant (which it is not). We did not preregister this analysis, because we did not expect this valuation to be so high.

[10] A reviewer also suggested that the certain discount may be considered weird in these studies.

References

Andreoni J, Sprenger C (2011) Uncertainty equivalents: Testing the limits of the independence axiom. NBER Working Paper 17342, National Bureau of Economic Research, Cambridge, MA.

Ariely D, Loewenstein G, Prelec D (2006) Tom Sawyer and the construction of value. *J. Econom. Behav. Organ.* 60(1):1–10.

Chow CC, Sarin RK (2001) Comparative ignorance and the Ellsberg paradox. *J. Risk Uncertainty* 22(2):129–139.

Ellsberg D (1961) Risk, ambiguity, and the Savage axioms. *Quart. J. Econom.* 75(4):643–669.

Fox CR, Tversky A (1995) Ambiguity aversion and comparative ignorance. *Quart. J. Econom.* 110(3):585–603.

Fox CR, Weber M (2002) Ambiguity aversion, comparative ignorance, and decision context. *Organ. Behav. Human Decision Processes* 88(1):476–498.

Frisch D, Baron J (1988) Ambiguity and rationality. *J. Behavioral Decision Making* 1(3):149–157.

Gneezy U, List J, Wu G (2006) The uncertainty effect: When a risky prospect is valued less than its worst possible outcome. *Quart. J. Econom.* 121(4):1283–1309.

Goldsmith K, Amir O (2010) Can uncertainty improve promotions? *J. Marketing Res.* 47(6):1070–1077.

Jung MH, Perfecto H, Nelson LD (2016) Anchoring in payment: Evaluating a judgmental heuristic in field experimental settings. *J. Marketing Res.* 53(3):354–368.

Keren G, Gerritsen LEM (1999) On the robustness and possible accounts of ambiguity aversion. *Acta Psychologica* 103(1–2): 149–172.

Keren G, Willemsen MC (2009) Decision anomalies, experimenter assumptions, and participants' comprehension: Revaluating the uncertainty effect. *J. Behav. Decision Making* 22(3):301–317.

Lichtenstein S, Slovic P (2006) *The Construction of Preference* (Cambridge University Press, Cambridge, UK).

Mazar N, Shampanier K, Ariely D (2017) When retailing and Las Vegas meet: Probabilistic free price promotions. *Management Sci.* 63(1):250–266.

McGraw AP, Shafir E, Todorov A (2010) Valuing money and things: Why a $20 item can be worth more and less than $20. *Management Sci.* 56(5):816–830.

Moon A, Nelson LD (2017) The uncertain value of uncertainty: When consumers are unwilling to pay for what they like. (March 1), https://ssrn.com/abstract=2676699.

Newman GE, Mochon D (2012) Why are lotteries valued less? Multiple tests of a direct risk-aversion mechanism. *Judgment Decision Making* 7(1):19–24.

Rottenstreich Y, Hsee CK (2001) Money, kisses, and electric shocks: On the affective psychology of risk. *Psych. Sci.* 12(3):185–190.

Rydval O, Ortmann A, Prokosheva S, Hertwig R (2009) How certain is the uncertainty effect? *Experiment. Econom.* 12(4):473–487.

Simonsohn U (2009) Direct risk aversion: Evidence from risky prospects valued below their worst outcome. *Psych. Sci.* 20(6): 686–692.

Wang Y, Feng T, Keller LR (2013) A further exploration of the uncertainty effect. *J. Risk Uncertainty* 47(3):291–310.

Yang Y, Vosgerau J, Loewenstein G (2013) Framing influences willingness to pay but not willingness to accept. *J. Marketing Res.* 50(6):725–738.