Research Article

Direct Risk Aversion

Evidence From Risky Prospects Valued Below Their Worst Outcome

Uri Simonsohn

University of California, San Diego

ABSTRACT—Why would people pay more for a $50 gift certificate than for the opportunity to receive a gift certificate worth either $50 or $100, with equal probability? This article examines three possible mechanisms for this recently documented uncertainty effect (UE): First, awareness of the better outcome may devalue the worse one. Second, the UE may have arisen in the original demonstration of this effect because participants misunderstood the instructions. Third, the UE may be due to direct risk aversion, that is, actual distaste for uncertainty. In Experiment 1, the UE was observed even though participants in the certainty condition were also aware of the better outcome; this result eliminates the first explanation. Experiment 2 shows that most participants understand the instructions used in the original study and that the UE is not caused by the few who do not. Overall, the experiments demonstrate that the UE is robust, large (prospects are valued at 65% of the value of the worse outcome), and widespread (at least 62% of participants exhibit it).

Why are people risk averse? This question has received considerable attention on the part of decision-making researchers, both theoretical and empirical, going back at least to Bernoulli (1738/1954). Despite the great number of theories that have been proposed to explain risk aversion, it is striking that the notion that people simply dislike uncertainty—that is, that uncertainty itself influences utility—is not part of any mainstream theory.1

Widely accepted risk-aversion theories, including expected-utility and prospect theory, arrive at risk aversion only indirectly, as a side effect of how outcomes are valued or how probabilities are judged. According to expected-utility theory, for instance, people are risk averse because they are satiated through consumption, and hence potential increases in wealth are valued less than potential decreases. Despite important conceptual differences between expected-utility and prospect theory, they share (together with all mainstream theories of risk aversion) a reliance on “indirect risk aversion.” All such theories make the following falsifiable prediction: An individual cannot be so risk averse as to value a risky prospect less than the prospect’s worst possible outcome.

A recent article by Gneezy, List, and Wu (2006), however, reported evidence that contradicted this consensus prediction. In several between-participants studies, Gneezy et al. found that people are willing to pay less, on average, for a binary lottery than for its worse outcome, a finding they coined the uncertainty effect (UE). For example, they found that people are willing to pay an average of $26 for a $50 gift certificate from Barnes and Noble, but only $16 for a lottery that pays either a $50 or a $100 gift certificate, with equal probability.

Given the potential importance of the UE, which poses a direct challenge to the overarching paradigm currently used to understand risk aversion, I set out to distinguish among three possible causes behind the UE. The first and most interesting possibility is that uncertainty enters directly into people’s utility function—that people exhibit what I refer to as direct risk aversion. The other two possibilities, which I discuss in some detail, are (a) that awareness of the lottery’s high-value outcome diminished the perceived value of the low-value outcome and hence of the lottery as a whole, and (b) that respondents erroneously believed the lottery could result in a payment of $0.

In all the demonstrations of the UE by Gneezy et al. (2006), participants evaluating lotteries were shown the two outcomes for each, whereas those evaluating outcomes to be received with certainty were shown just one. This means that the manipulation


1A small theoretical literature has considered the possibility of direct utility from risk, primarily to account for gambling (see, e.g., Conlisk, 1993; Diecidue, Schmidt, & Wakker, 2004; Fishburn, 1980; Royden, Suppes, & Walsh, 1959).
of uncertainty was fully confounded with the number of outcomes presented to participants. As it turns out, this aspect of the design could—independently of uncertainty—explain the UE. Abundant research has examined the differences between evaluations performed on single versus multiple items (e.g., Bazerman, Loewenstein, & White, 1992; Bazerman, Moore, Tenbrunsel, Wade-Benzoni, & Blount, 1999; Hsee, 1996, 1998; Hsee & Zhang, 2004; List, 2002; for a review, see Hsee, Loewenstein, Blount, & Bazerman, 1999). For present purposes, a particularly relevant finding in this literature is that the willingness to pay (WTP) for low-quality items drops when they are evaluated jointly with similar but superior ones. For example, Hsee (1996) reported that participants were willing to pay $24 for a dictionary with 10,000 words when they were evaluating just that dictionary, but only $19 for the same dictionary if they were also evaluating a dictionary with 20,000 words.

The fact that a lottery includes a similar yet superior outcome, therefore, may diminish the perceived value of the worse outcome. This mechanism provides a plausible uncertainty-independent explanation for the UE. Some patterns in the data obtained by Gneezy et al. (2006) suggest that the use of separate evaluation played a role in how outcomes to be received with certainty were valued. For example, the median WTP to receive $100 in a year was the same as the median WTP to receive $200 in a year: $50 (see Table I in Gneezy et al.); such inadequate sensitivity to changes in the quantity of the good being evaluated is typical of separate (but not joint) evaluations.

The experiments presented in this article eliminated this confound by asking participants in the certainty condition to value both the low- and the high-value outcome. That is, participants in both conditions conducted joint evaluations.

Another possible cause underlying the findings of Gneezy et al. (2006) is that the lottery descriptions they employed did not unambiguously rule out a $0 payment. One of their lottery descriptions, for instance, was “a lottery that pays $50 or $100 with equal probability” (p. 1304). Although the authors were referring to a lottery that would always pay one of two outcomes, this description is also consistent with a lottery that could result in a payment of $0, such as one paying $0 with 98% probability and paying $50 or $100 each with 1% probability.2

The possibility that the findings of Gneezy et al. (2006) were caused by participants’ misunderstanding of the lottery has been put forward by two independent sets of researchers (Keren & Willemesen, in press; Ortmann, Prokosheva, Rydval, & Hertwig, 2007). To address this potential explanation, I modified the lottery description in Experiment 1 to more definitely rule out a $0 payment and included comprehension questions about the lottery in Experiment 2.

EXPERIMENT 1: THE UNCERTAINTY EFFECT WITH JOINT EVALUATION

The design of Experiment 1 incorporated two notable modifications of the experiments conducted by Gneezy et al. (2006): First, participants in the certainty condition indicated their WTP for two outcomes rather than a single outcome. Second, the lottery descriptions were modified slightly, in an attempt to eliminate the possibility that participants would believe a $0 payment was possible.

Method

Two hundred seventy-nine participants, primarily University of Pennsylvania undergraduates, answered hypothetical WTP questions as part of a series of surveys and experiments they completed in exchange for monetary payment at Wharton’s Behavioral Lab.

The experiment had a 6 × 2 between-participants design, in which six conditions, each presenting a different pair of low- and high-value items, were crossed with elicitation mode, that is, whether participants indicated their WTP for each of the two items (certainty condition) or for the corresponding 50:50 lottery (uncertainty condition). Each participant evaluated only one item pair under one elicitation mode. The six item pairs were combinations of the following: $50 and $100 gift certificates for Barnes and Noble; $50 and $100 gift certificates for Pod, a high-end Asian fusion restaurant located near the University of Pennsylvania; and free three-course meals, also at Pod, for either two or four people (see Table 1).

A fundamental feature of the design is that participants in the certainty condition were aware of both items before indicating their WTP for the first one (if that were not the case, the first elicitation would not have been performed under joint evaluation). This was accomplished with the following instructions:

We are interested in how much you would be willing to pay for two different items. In particular we will ask you how much you would be willing to pay for [low-value item] and for [high-value item].

If you could only buy the [low-value item], what is the highest amount of money you would pay for it?____

If you could only buy the [high-value item], what is the highest amount of money you would pay for it?____

Also, as mentioned earlier, the format of the lottery descriptions was modified to reduce ambiguity as to the possibility of a $0 payment. Specifically, the descriptions stated:

We are interested in how much you would be willing to pay for a lottery ticket that will for sure pay one of two possible rewards (both are equally likely).

It will either pay [low-value item] or [high-value item].

---

2In some experiments, Gneezy et al. used the description “a lottery ticket that gives you a 50 percent chance at a $50 gift certificate, and a 50 percent chance at a $100 gift certificate.” This could be interpreted as consisting of two independent 50:50 draws, one for $50 and one for $100, which would make a $0 (and a $150) outcome possible.
What is the highest amount of money you would pay for this lottery?___

This study was part of a series of experiments completed by participants in individual cubicles with personal computers. Participants were assigned to conditions sequentially, according to the order in which they logged on to the survey.

**Results**

Table 1 summarizes the results for Experiment 1. The UE was observed for both the average and the median WTP for all six item pairs. Four of the six differences in means were statistically significant at the .05 level, and the *p* values for the remaining two values were slightly greater than .10. Four of the six Wilcoxon tests were significant at the .05 level, and the remaining two were significant at the .10 level.

The ratio between the WTP for the lottery and the WTP for the low-value item gives an indication of how large the UE was. For the means, the average ratio across the six item pairs was .65, indicating that participants were willing to pay two thirds as much for the lotteries as for the lotteries’ lowest possible outcomes. For the medians, the average ratio was a striking .38.

Figure 1 plots the cumulative distribution of the WTP for the $50 Barnes and Noble gift certificate and for the 50:50 lottery for a $50 or $100 gift certificate at the same store. The figure shows, for instance, that for around 50% of participants, the WTP for the lottery was less than $20, whereas not a single participant’s WTP for the $50 gift certificate was lower than $20; in other words, around half the participants in the uncertainty condition valued the lottery less than anyone in the certainty condition valued its worse outcome. This pattern suggests that the uncertainty condition valued the lottery less than anyone in the certainty condition valued its worse outcome. This pattern suggests that the uncertainty condition valued the lottery less than anyone in the certainty condition valued its worse outcome.

Although central-tendency measures in Experiment 1 did yield a UE, it is of interest to consider how widespread the UE was across participants. This calculation is not entirely straightforward because the design was between participants, and hence only one of the two relevant valuations was observed for any given participant (i.e., one cannot simply count the number of participants giving a higher WTP for the low-value item than for the lottery). One can, nevertheless, learn quite a bit by comparing the distributions of valuations across conditions.

**Note.** The *t*-test and Wilcoxon-test columns report *p* values from tests comparing willingness to pay for the low outcome and willingness to pay for the lottery.

### Table 1

<table>
<thead>
<tr>
<th>Item pair</th>
<th>Willingness to pay ($)</th>
<th>High outcome</th>
<th>Low outcome</th>
<th>Lottery</th>
<th>Wilcoxon test</th>
<th>Certainty condition</th>
<th>Uncertainty condition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High outcome</td>
<td>Low outcome</td>
<td>Lottery</td>
<td>Wilcoxon test</td>
<td>Certainty</td>
<td>Uncertainty</td>
<td></td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>Mdn</td>
<td>M</td>
<td>Mdn</td>
<td>M</td>
<td>Mdn</td>
<td>t test</td>
</tr>
<tr>
<td>Bookstore: $100 certificate</td>
<td>84.2</td>
<td>90</td>
<td>43.2</td>
<td>47</td>
<td>23.2</td>
<td>12.5</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Restaurant: $100 certificate</td>
<td>67.6</td>
<td>65</td>
<td>34.5</td>
<td>30</td>
<td>25.7</td>
<td>15</td>
<td>.133</td>
</tr>
<tr>
<td>Bookstore: $100 certificate</td>
<td>57.8</td>
<td>60</td>
<td>33.7</td>
<td>35</td>
<td>23.9</td>
<td>10</td>
<td>.104</td>
</tr>
<tr>
<td>Restaurant: $100 certificate</td>
<td>73.5</td>
<td>80</td>
<td>40.8</td>
<td>45</td>
<td>25.1</td>
<td>20</td>
<td>.066</td>
</tr>
<tr>
<td>Restaurant: free meal for four</td>
<td>91.0</td>
<td>100</td>
<td>35.6</td>
<td>40</td>
<td>24.5</td>
<td>12.5</td>
<td>.044</td>
</tr>
<tr>
<td>Restaurant: free meal for four</td>
<td>97.7</td>
<td>85</td>
<td>46.9</td>
<td>42.5</td>
<td>30.7</td>
<td>20</td>
<td>.016</td>
</tr>
</tbody>
</table>

*Note. The *t*-test and Wilcoxon-test columns report *p* values from tests comparing willingness to pay for the low outcome and willingness to pay for the lottery.
that the UE is quite prevalent, and not driven by a few extreme responses.

Employing the distributions depicted in Figure 1, one can calculate the precise share of participants who exhibited the UE, if one assumes that the relative ranking of valuations across participants is constant across conditions, that is, if one assumes that if a participant would pay more for the low-value item than another participant does, then he or she would also pay more for the lottery. This would mean that the person paying the most for the lottery is assumed to be the person paying the most for the low-value item, the one paying the second-highest amount for the lottery is assumed to be the second-highest payer for the low-value item, and so on. Under this assumption, the point at which the two distributions cross indicates the share of people who are willing to pay less for the lottery than for the low-value item. In Figure 1, the cumulative distributions cross at roughly 80%, which suggests that around 80% of participants exhibited the UE. The crossing points for the other five item pairs ranged between 70% and 80%. (Readers can view figures showing cumulative distributions for these other pairs in the Supporting Information available on-line; see p. 692.)

Because the assumption on which these estimates rest cannot be tested, I also estimated the lowest possible share of participants exhibiting the UE. To this end, I matched participants between conditions in a manner that minimized the total number of “violations.” For example, imagine a situation in which there were 3 participants per condition, with the valuations of the lottery being $10, $20, and $30, and the valuations of the low-value item being $15, $25, and $35. For expository purposes, let us begin again with the assumption that the ranking of valuations is constant across conditions. Under this assumption, 100% of the participants exhibit the UE (because the lowest, middle, and highest valuations of the lottery are lower than the corresponding lowest, middle, and highest valuations of the low-value item). In contrast, the lowest possible share of participants exhibiting the UE is just one third. If the participant paying $35 for the low-value outcome were the one paying $10 for the lottery, and the participants paying $15 and $25 for the low-value outcome were willing to pay $20 and $30 for the lottery, then only 1 participant would exhibit the UE.

I conducted analogous calculations for the distributions of valuations for all six item pairs. The lower bound on the share of participants exhibiting the UE ranged between 52% and 70%, with an overall mean of 62%. In sum, if we assume that the ranking of valuations of the lottery and the ranking of valuations of the low-value item are the same across subjects between conditions, then the best estimate of the share of participants exhibiting the UE is about 80%. If we are not willing to make this assumption, we can place the lower bound on this share at 62%.

The main findings of Experiment 1, then, are that the UE (a) is not caused (at least exclusively) by the fact that the lottery contains a superior outcome that may reduce the valuation of the low-value one, (b) is large in magnitude, and (c) is widespread across participants. Experiment 2 addressed the possibility that people value the binary lottery less than its actual worse outcome not because of direct risk aversion, but because they incorrectly believe that the lottery may result in a $0 payoff.

**EXPERIMENT 2: DO PARTICIPANTS UNDERSTAND THE LOTTERY?**

If participants in UE studies erroneously believe that, as in most lotteries, a $0 payoff is possible, it would not be surprising that they would pay less for the lottery than for the low-value outcome. Keren and Willemesen (in press) and Ortmann et al. (2007) proposed that this is all that is behind the findings of Gneezy et al. (2006). Both studies assessed how modifying instructions and the randomizing procedure behind the lottery affected the prevalence of the UE. They both found that these modifications do (sometimes) attenuate and even eliminate the UE for the average valuation (but recall the previous discussion regarding the limitations of inferences about the UE based on mean and median valuations).

Keren and Willemesen (in press) also found that a substantial share of participants erroneously answer a comprehension question about the lottery and that the UE is observed only for these participants. In particular, they asked: “True/False: The lottery offered me, with 100% certainty, at least 50 euros in Book certificates.” Only for the roughly 45% (on average, across experiments) of participants who answered “false” was WTP for the lottery lower than WTP for the low-value outcome. These results are consistent with the UE being driven by misunderstood instructions, rather than by a direct distaste for uncertainty. Keren and Willemesen’s design and analysis, however, contained two potential problems. First, participants may have found the comprehension question itself difficult to comprehend. Second, only participants in the uncertainty condition answered the comprehension question (and hence could potentially be excluded), and selective elimination of observations in only one of the two conditions being compared may have introduced bias. Hence, it is possible that, despite Keren and Willemesen’s findings, the UE is fully caused by a direct distaste for uncertainty.

In light of these considerations, Experiment 2 contained two alternative comprehension questions. In one condition, participants were directly asked to state the lowest possible payoff of the lottery. Answers to this question provided a simple assessment of the share of participants who erroneously believed they could earn less than what the lottery actually could pay. In the other condition, participants were instead asked the true/false comprehension question. In addition, all participants indicated their WTP both for an item to be obtained with certainty and for a lottery. By comparing the WTP for the sure-thing item across participants who answered the comprehension questions cor-
correctly and incorrectly, I estimated the statistical bias introduced by dropping observations only from the uncertainty condition.

Method
One hundred ninety-six participants, primarily University of Pennsylvania undergraduates, participated in this study as part of a series of surveys and experiments. The experiment had a simple two-condition between-participants design, with the comprehension question systematically manipulated between participants. All questions concerned hypothetical scenarios.

All participants began by indicating their WTP, that day, for $100 to be received in a year. They then all indicated their WTP for a 50:50 lottery that paid either a $50 Barnes and Noble gift certificate or a $100 gift certificate at the same store (the lottery description from Experiment 1 was used). Participants were then asked one of two comprehension questions. Half the participants were asked: “True/False: The lottery from the previous question offered me, with 100% certainty, at least $50 in gift certificates.” The other half were asked: “What was the lowest possible payment the lottery could pay?” These participants were given a multiple-choice set of answers that included gift-certificate values from $0 to $250, in steps of $25. Note that participants could no longer see the lottery description when they were asked the comprehension question.

Results
Table 2 summarizes the results from Experiment 2. Arguably the most important findings are that of the 97 participants assigned to the multiple-choice question, 92% (n = 89) correctly answered that $50 was the lowest possible payoff, and that not a single participant answered 80, the most relevant value associated with misunderstanding. Two of the 8 participants who answered incorrectly indicated that the lowest possible payment was $25, whereas the other 6 responded with a number greater than $50. If anything, these results suggest that participants who misunderstand the lottery overestimate its worse outcome; more likely, rather than signaling lack of understanding of the lottery per se, answering the comprehension question incorrectly signals not having taken the task seriously enough.

Nevertheless, it is worthwhile assessing the extent to which participants such as these may be behind the UE. Table 2 shows that the WTP for the lottery was in fact much lower among the 8 participants who gave an erroneous answer to this comprehension question (M = 15.5) than among the 89 who answered it correctly (M = 31.1). It is tempting to interpret this result as suggesting that misunderstood instructions do indeed play a role in the UE. Such an interpretation, however, relies on the assumption that participants who answer the comprehension question correctly and incorrectly do not differ in their (reported) valuation of items to be received for certain. This assumption is contradicted by the fact that participants who gave an incorrect answer to this comprehension question also reported a dramatically lower WTP for $100 to be received in a year than did participants who answered the question correctly. This pattern was found for both mean and median valuations. This finding exemplifies the problems associated with comparing a subset of participants in one treatment with the full set of participants in another.

As was the case for the multiple-choice question, participants who answered the true/false question erroneously gave a lower

<table>
<thead>
<tr>
<th>Comprehension question</th>
<th>Receiving $100 in a year</th>
<th>Lottery for a $50 or $100 gift certificate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean WTP ($)</td>
<td>Median WTP ($)</td>
</tr>
<tr>
<td>Multiple-choice</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Answered correctly (n = 89)</td>
<td>57.98</td>
<td>50</td>
</tr>
<tr>
<td>Answered incorrectly (n = 8)</td>
<td>15.75</td>
<td>10</td>
</tr>
<tr>
<td>Total (n = 97)</td>
<td>54.50</td>
<td>50</td>
</tr>
<tr>
<td>True/false</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Answered correctly (n = 78)</td>
<td>47.38</td>
<td>50</td>
</tr>
<tr>
<td>Answered incorrectly (n = 21)</td>
<td>52.43</td>
<td>30</td>
</tr>
<tr>
<td>Total (n = 99)</td>
<td>48.45</td>
<td>50</td>
</tr>
</tbody>
</table>
median WTP for $100 in a year than did participants who answered this question correctly. The mean WTPs did not show this pattern because 2 participants who answered the true/false question incorrectly reported WTPs of $200. If only participants with WTPs of $100 or less are considered (in both conditions), the pattern seen in the medians is present: Average WTP for $100 in a year was lower among participants who answered the true/false question incorrectly (M = 33.11) than among those who answered it correctly (M = 45.34).

Experiment 2, in sum, suggests that the mechanism behind the UE is not the erroneous belief that the lotteries may pay a counterfactually low outcome. Experiment 2 also highlights the importance of conducting symmetric comprehension checks across conditions in order to avoid statistical bias caused by, for example, eliminating only from one condition the participants who do not take the task seriously.

**GENERAL DISCUSSION**

The results from these experiments suggest that the UE, valuing a risky prospect below the value of its worse possible outcome, occurs as the consequence of direct risk aversion, that is, risk aversion that arises directly from a literal distaste for uncertainty, rather than indirectly as a consequence of how people value outcomes or weight probabilities. The UE was found to be quite large, with the average valuation of a lottery with two possible outcomes being just 65% of the average valuation of its lower outcome. The UE was also found to be widespread, with the lower bound on the percentage of participants exhibiting it being 62%. Although the notion of direct risk aversion has been considered before (within frameworks regarding the utility from gambling—see footnote 1), it has not entered mainstream understanding of risk-averse behavior. The robustness, magnitude, and prevalence of the UE suggest that this may be an important shortcoming in the current understanding of risky choice.

Future work on direct risk aversion should be aimed at making progress in two different directions. First, research should explore the ability of direct risk aversion to account for real-world behaviors that are at odds with existing theories. For example, expected-utility and prospect theory predict that people should not purchase insurance for small-stakes risks, yet such forms of insurance (e.g., warranties for electronics, insurance policies with low deductibles, and mail insurance) are popular. Direct risk aversion may explain why. The challenge, of course, is to avoid the tautological argument of alluding to direct risk aversion to “explain” otherwise puzzling risk-averse behavior. This leads to the second direction for future research on direct risk aversion: clarifying what determines the perceived riskiness of, and hence aversion to, a given prospect. It seems likely that factors that have previously been thought to influence risky choice independently—such as familiarity (Song & Schwarz, 2008; Weber, Siebenmorgen, & Weber, 2005), emotions (Loewenstein, Weber, Hsee, & Welch, 2001), and ambiguity aversion (Ellsberg, 1961; Fox & Tversky, 1995)—all ultimately influence risky choice by affecting the subjective sense of riskiness of a prospect, which in turn influences decisions via direct risk aversion. Research exploring this and related issues could reshape our understanding of decision making under uncertainty.

**REFERENCES**


Ottmann, A., Prokosheva, A., Rydval, O., & Hertwig, R. (2007). Valuing a risky prospect less than its worst outcome: Uncertainty...
**SUPPORTING INFORMATION**

Additional Supporting Information may be found in the on-line version of this article:

- **Figure S1**
- **Figure S2**
- **Figure S3**
- **Figure S4**
- **Figure S5**

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.