

# WHEN ARE DECISIONS IMPROVABLE? AN EVALUATION OF DIAGNOSTIC METHODS. DECISION CONFIDENCE METHOD FOLLOW-UP EXPERIMENT

B. DOUGLAS BERNHEIM, ALDO LUCIA, KIRBY NIELSEN, CHARLES D. SPRENGER

## 1. MOTIVATION

In the main experiment described in the original PAP, we evaluate three methods for identifying improvable choices: documenting specific misconceptions (the Characterization Assessment method), gauging confidence in choices (the Decision Confidence method), and showing that specific behavioral patterns in a domain of interest also emerge in a related domain where they are objectively suboptimal (the Pattern Matching method). Characterization Assessment is informative but narrow, which motivates consideration of the two broader methods. Results based on the Decision Confidence and Pattern Matching methods have conflicting implications: While the Pattern Matching method attributes small-stakes risk aversion and the Common Ratio Effect to improvability, the Decision Confidence method attributes the *opposite* patterns to improvability. In principle, the two methods may detect different types of errors. Alternatively, either or both of these methods may have methodological shortcomings. This conflict motivates a closer investigation of the assumptions underlying both methods. This document presents the pre-analysis plan for a follow-up experiment focusing on the Decision Confidence method.

The Decision Confidence method assumes that expressions of low confidence in choices reliably capture perceptions of improvability. The validity of this premise hinges on the assumption that respondents interpret the confidence question to mean precisely what the analyst intends. Whether this assumption is satisfied in the context of our tasks is an open empirical question. Less-than-complete confidence could indicate that a subject thinks they could have made a better decision based on what they knew or could have known when they made their choice. Alternatively, it could reflect uncertainty about how the choice will turn out, the possibility of learning new information, the perception that there are multiple valid views about what the subject wants, or the belief that the choice is close to a toss-up. These interpretations have different implications for perceived improvability. It is also an open question whether respondents interpret the same general confidence question similarly across Risky and Risk-Free domains. Accordingly, the confidence follow-up experiment evaluates whether a broad question about decision confidence provides a reliable diagnostic of perceived improvability.

## 2. CONFIDENCE AND RECLASSIFICATION FOLLOW-UP

**2.1. Experimental Design.** This experiment retains the Risky and Risk-Free choice tasks, confidence elicitation, and lottery-payment mechanism described in the original PAP. We do not elicit characterization tables. In the current protocol, subjects receive \$15 for completing the study, in addition to any earned quiz and lottery bonuses.

Subjects first complete one Risky and one Risk-Free practice task. They then receive instructions about the lottery-payment mechanism and answer three comprehension questions using a worked example: how the potential bonus payment is determined; what the potential bonus is if they choose Option B and a purple ticket is drawn; and under what condition they receive the potential bonus, including their 1-in-5 chance of eligibility. Incorrect answers receive explanatory feedback, and subjects must answer all three questions correctly before proceeding.

Each subject then completes 20 experimental tasks drawn from the same bank of 40 tasks. The bank contains ten tasks in each of four categories: Risky Unmixed, Risky Mixed, Risk-Free Unmixed, and Risk-Free Mixed. We randomly select five tasks without replacement from each category and present the resulting 20 tasks in random order. In every practice and experimental task, subjects first choose between Options A and B. We then ask: “On a scale from 0 (not confident at all) to 100 (completely confident), how confident do you feel about this choice?” Subjects respond using a slider.

Subjects who report a confidence level of 100 for all 20 choices skip the remainder of the confidence follow-up and proceed to payment. For subjects who report less-than-complete confidence at least once, we next explain the possible reasons for such a response. Drawing inferences about the nature of less-than-complete confidence from subjective responses raises a conceptual difficulty: Just as subjects may misinterpret questions about decision confidence, they may also misinterpret questions about the reasons for less-than-complete confidence. Carefully wording a definition in terms that are clear to an economist does not necessarily make it clear or memorable to an

experimental subject. Accordingly, we teach the relevant categories before asking subjects to apply them to their own choices. We organize the possible meanings of less-than-complete confidence into four General Reasons:

- (1) *Worry that one could have made a better decision without access to additional information.* This reason captures the possibility that a subject believes they did not correctly assess their options based on what they knew or could have known when they chose. We distinguish four possibilities: uncertainty about what the options mean, insufficient clarification of one’s objective, use of a poor mental shortcut, and incorrect application of one’s decision criterion.
- (2) *Learning more after making the decision.* This reason captures the possibility that later information may change how the subject evaluates the choice, without implying that they could have chosen better based on what they knew or could have known at the time. We distinguish 20–20 hindsight from uncertainty about how good the chosen option will turn out to be even when it is best regardless.
- (3) *Multiple valid views about what one wants.* This reason captures the possibility that the subject sees more than one valid way to evaluate their objectives.
- (4) *A toss-up.* This reason captures the possibility that the subject views the two options as nearly equally desirable.

We also include *Other Reasons* as a residual category. Our description of each General Reason has three parts: a conceptual explanation, a prototypical example, and a memorable label. To avoid priming and experimenter demand effects, the examples come from another choice domain: a person named Chris choosing between social events. For the first two General Reasons, we separately explain and illustrate each of the possibilities listed above. Other Reasons has two non-exhaustive examples: Chris cannot know, either before or after choosing, whether a dinner-party host genuinely wants him to attend; and Chris feels uneasy about committing to a baseball game despite believing it is clearly his best choice, because he habitually worries about decisions. We present all four General Reasons and Other Reasons before asking subjects to relate them to their own choices.

The numbering above and throughout this PAP identifies the meanings of the General Reasons, not their display positions. For each subject who enters the follow-up, we draw one order uniformly from the 24 permutations of the four General Reasons before the first explanation. We retain that order throughout the explanation blocks, both quizzes’ answer options and reference lists, the rating walkthrough, and the rating tasks. The explanations use the neutral heading “General Reason” together with each reason’s descriptive title. Other Reasons is presented last in the explanations and ratings. The questions within each quiz remain in the fixed sequences reported in Appendix A.

After presenting the four General Reasons and Other Reasons, we administer an incentivized classification test consisting of seven one-shot questions involving decisions in domains unrelated to the experiment. Each question describes why a person expresses less-than-complete confidence and asks the subject to identify the General Reason that best fits the description. Subjects choose among the four General Reasons; Other Reasons is not an answer option. A subject earns \$0.25 for every correct answer, up to \$1.75 for the seven questions. After the test, we report the subject’s score and provide item-by-item feedback. Appendix A reports the exact wording and correct answers for this test.

After subjects receive this feedback, an interactive walkthrough introduces the rating table and its summary. Subjects can inspect a filled-out example and modify its ratings to see how the summary changes; the example ratings are explicitly described as demonstrations rather than suggested answers. We then present up to eight earlier tasks for the follow-up questions. These tasks are selected immediately after the twentieth confidence response, before the explanations and General-Reason quiz. Only tasks with confidence below 100 are eligible. We initially select up to two lowest-confidence eligible tasks from each of the four categories. When a category contains fewer than two eligible tasks, we fill the shortfall, when possible, using the next-lowest-confidence eligible tasks not already selected from the other Mixed or Unmixed category within the same domain. Thus, a missing Risky task can be replaced only by another Risky task, and a missing Risk-Free task only by another Risk-Free task. We randomize both ties in confidence and the order of the selected tasks. For each selected task, we show the original task, the subject’s choice, and their confidence rating. Subjects then rate how much each of the four General Reasons, as well as Other Reasons, accounted for their less-than-complete confidence on a scale from 1 (Very Little) to 7 (Very Much). The table uses the subject’s previously assigned reason order, with Other Reasons last. Definitions and examples remain available through popups and expandable references in the walkthrough and rating tables.

After subjects complete these ratings, we administer a second incentivized classification test as an additional validation check. The first classification test assesses whether subjects can distinguish the General Reasons in unrelated choice domains; the second assesses whether they interpret those reasons as intended in the context of the experimental choices.

For this test, we independently select one task uniformly at random from five Risky Unmixed tasks in which Option A provides a certain payoff and Option B has positive payoff variance and an expected value \$1 higher.

These tasks present an economically meaningful tradeoff between expected value and payoff variability. We use the same selected task in all nine questions, holding the choice problem fixed while varying the explanation for less-than-complete confidence. Each question describes a different person who chose Option B and expressed less-than-complete confidence, then asks the subject to identify the General Reason that best fits the person’s explanation. Subjects again choose among the four General Reasons. A subject earns \$0.25 for every correct answer, up to \$2.25 for the nine questions. We report the subject’s score and provide item-by-item feedback after the test. Appendix A reports the exact wording and correct answers for this test. The two understanding bonuses are independent of one another and of lottery eligibility, so subjects who enter the follow-up can earn up to \$4.00 from the classification tests. Payment-comprehension questions are unpaid.

**2.2. Analysis.** The first step in our analysis will be to describe behavior and confidence in the full choice sample. These analyses reproduce the analyses of choice and confidence specified for the main experiment in the original PAP. We define the *full choice sample* as all 20 completed choice–confidence observations for every subject. We will:

- Document Safe/“Safe” choice rates and expected-value maximization rates by Risky versus Risk-Free, Mixed versus Unmixed, probability, and the +1 versus −1 version of a task. We will also calculate the magnitude of the Common Ratio Effect as defined in the original PAP.
- Relate confidence to Safe/“Safe” choice and expected-value maximization. Following the analysis specified in the original PAP, we will divide confidence into nine bins and, within each bin, calculate average confidence, the share choosing the Safe/“Safe” option, and the share maximizing expected value. We will present these relationships separately for Mixed and Unmixed tasks in each domain, report the corresponding correlations, and display 95% confidence intervals.

We next turn to subjects’ reported reasons for less-than-complete confidence. We define the *confidence rationalization sample* as the selected tasks for which subjects rate the General Reasons for their less-than-complete confidence. Subjects who report confidence of 100 for all 20 choices contribute to the full choice sample but, by design, not to the confidence rationalization sample. For each task in the confidence rationalization sample, we classify every category receiving the highest rating as a most important reason. Thus, in the event of a tie, more than one category is classified as most important and the resulting frequencies need not sum to 100%.

We aggregate the categories according to their implications for perceived improbability. General Reason 1, *Worry that one could have made a better decision without access to additional information*, implies perceived improbability. General Reasons 2 through 4 do not imply perceived improbability, and Other Reasons is an unclassified residual category. An observation can belong to more than one of these groups when categories from different groups are tied for the highest rating.

Our analysis of these rationales will proceed in two steps.

- (1) *Rationales for less-than-complete confidence.* We will report the frequency of less-than-complete confidence in the full choice sample and the number and composition of tasks in the confidence rationalization sample. We will then report how frequently each of the four General Reasons and Other Reasons receives the highest rating. We will also report how frequently General Reason 1 receives the highest rating, how frequently at least one of General Reasons 2 through 4 receives the highest rating, and how frequently Other Reasons receives the highest rating. These results will be presented primarily for Risky and Risk-Free choices and secondarily for each of the four Risky/Risk-Free by Mixed/Unmixed design cells. We will compare the frequencies of the perceived-improvability and non-perceived-improvability indicators within and across the Risky and Risk-Free domains.
- (2) *Classification tests.* We will analyze the two classification tests separately. The first tests whether subjects can classify the General Reasons in choice domains unrelated to the experiment and occurs before subjects rate the reasons for their own less-than-complete confidence. The second tests whether subjects can apply the General Reasons in the context of a Risky choice task from the experiment.
  - (a) *Classification test in domains unrelated to the experiment.* We will report the score distribution for the classification test in unrelated domains. Our primary rationale analysis will include all subjects in the confidence rationalization sample, regardless of test performance, and we will examine sensitivity of results to this classification test score.
  - (b) *Classification test in the context of the experiment.* We will report the score distribution for the classification test in the context of the experiment. Examples 1, 3, 6, and 9 involve perceived improbability because General Reason 1 is correct; examples 2, 4, 5, 7, and 8 do not because one of General Reasons 2 through 4 is correct. We define two directional classification errors: choosing General Reason 1 in the five non-improvability examples, and choosing one of General Reasons 2 through 4 in the four

improvability examples. For each subject, we will calculate the frequency of each directional error across the relevant examples. We will also report each error rate pooled across the relevant examples and separately by example. Our primary rationale analysis will include all subjects in the confidence rationalization sample, regardless of test performance, and we will examine sensitivity of results to this classification test score.

**2.2.1. Sample Size and Power Calculations.** We plan to recruit 450 subjects. Using data from the control condition of the main experiment, we evaluate the implications of this sample size for the Common Ratio Effect, the rationales for less-than-complete confidence, and the study-context classification test.

- (1) *Common Ratio Effect.* Following the definition in the original PAP, we calculate the Common Ratio Effect for each pair of associated Unmixed and Mixed tasks as the difference in the probability of choosing the safer option. We then average these differences across the ten task pairs in each domain. In the control condition of the main experiment, the resulting effect was 16.0 percentage points in the Risky domain and 13.1 percentage points in the Risk-Free domain.

Each subject receives five of the ten tasks in each design category, so 450 subjects provide 225 observations per task in expectation. As a conservative benchmark, we treat the observations in an associated Unmixed and Mixed task as independent and use the average observed choice probabilities in each domain. At the 5% significance level, a two-sided test then has approximately 95% power to detect the observed Risky effect and 81% power to detect the observed Risk-Free effect. This calculation does not credit the design with any additional precision from averaging across task pairs.

- (2) *Confidence rationalization sample.* Our principal analysis concerns the rationales for less-than-complete confidence in the Risky domain. To estimate the available sample under our balanced task design, we conduct 2,500 bootstrap replications. In each replication, we resample 450 control-condition subjects with replacement and, within each design category, draw five confidence reports with replacement before applying the experiment’s follow-up selection rule.<sup>1</sup> In the main experiment, confidence was below 100 in 72.6% of Risky choices and 52.7% of Risk-Free choices. The resampling procedure yields approximately 1,607 selected Risky tasks from 415 subjects and 1,251 selected Risk-Free tasks from 344 subjects.

Within each domain, we compare the frequency with which each General Reason is top-rated. Because ties can activate both indicators, we use the maximum possible variance of their difference. Under the baseline assumption that observations within subjects are independent, a two-sided test at the 5% significance level has 80% power to detect a difference of 7.0 percentage points in the Risky domain and 7.9 percentage points in the Risk-Free domain.<sup>2</sup> We will also compare the frequency with which each General Reason is top-rated across Risky and Risk-Free domains.

- (3) *Classification tests.* The resampling procedure predicts that approximately 425 subjects will report confidence below 100 at least once and complete both classification tests. For the study-context test, we compare how often each question is misclassified. Under the baseline assumption that the nine responses are independent, setting each error probability to 0.5 maximizes the variance. A two-sided test at the 5% significance level then has 80% power to detect a difference of 4.6 percentage points.<sup>3</sup>

### 3. FURTHER EXPERIMENTS ON ROBUSTNESS

Our selection of our general confidence question follows from a principled evaluation of confidence questions in the prior literature. These confidence questions vary in several dimensions including whether they include a reference to a “best” selection, and a reference to conditioning statements such as “given your preferences and information”. For the purposes of investigating improvability, there are substantial challenges to including these features, but it is useful to assess the robustness of our results to these wording adjustments. To this end, we propose to repeat the analysis outlined above with confidence statements altered to refer to a best selection conditional on preferences and information.

<sup>1</sup>One control subject has no response in one design category. For that subject-category, we use the pooled confidence responses from the same category and condition.

<sup>2</sup>With a within-subject correlation of 0.5, the minimum detectable differences are 11.0 percentage points in the Risky domain and 12.3 in the Risk-Free domain; with a correlation of 1, they are 13.9 and 15.5 percentage points, respectively.

<sup>3</sup>Allowing arbitrary dependence among a subject’s nine responses and using the maximum possible variance of their difference raises the minimum detectable difference to 13.6 percentage points.

## APPENDIX A. CONFIDENCE RECLASSIFICATION QUIZ QUESTIONS

This appendix reports the exact scenarios and prompts used in the two classification tests, along with their correct answers. Questions appear in the fixed sequences shown below. Every question uses the following answer options, numbered here by meaning rather than display position. Their on-screen order follows the subject’s single randomized order, shared with the explanations, reference lists, rating walkthrough, and rating tasks:

- (1) You Worry You Could Have Made a Better Decision Without Access to Additional Information

*Possibilities include (but are not necessarily limited to):*

- Weren’t Sure What Options Mean
- Might Not Have Fully Clarified Your Objective
- Might Have Used A Poor Shortcut
- Might Not Have Applied Your Decision Criterion Correctly

- (2) You Will Learn More After Making Your Decision

*Possibilities include (but are not necessarily limited to):*

- 20-20 Hindsight
- Not Sure How Good My Chosen Option Will Be, Even If It’s Best Regardless

- (3) Multiple Valid Views About What You Want

- (4) It’s a Toss-Up

*Other Reasons* is not an answer option in either classification test.

## A.1. General-Context Classification Test.

**Understanding Question 1 of 7.**

Susan is deciding between two mobile phones. The phones have dozens of features that differ; when she makes her choice, she is aware of all of them, and accurate information about all of them is easily available. When she focuses on evaluating a feature, she has no trouble visualizing accurately how she’ll end up feeling about it. However, she doesn’t take the time to focus on evaluating all of them because she has other things to do. Instead, she streamlines her decision process by focusing only on the two that are most important to her: screen size and battery life. She reports overall confidence less than 100 because she wonders whether she might have changed her choice if she had also taken the time to focus on evaluating the other features.

*Question:* Which of the following general reasons best explains Susan’s less-than-complete confidence?

**Correct answer:** You Worry You Could Have Made a Better Decision Without Access to Additional Information.

**Understanding Question 2 of 7.**

Edward is deciding between two options. Option A is a box of oatmeal cookies, while Option B is a box containing either dark chocolates or milk chocolates. For Option B, he will not know whether the box contains dark chocolates or milk chocolates unless he chooses that option and opens the box, at which point it will be too late to change his mind. If he knew the box contained dark chocolates, he would definitely prefer Option B, but if he knew the box contained milk chocolates, he would definitely prefer Option A. He selects Option B, but he reports overall confidence less than 100 because once he sees what’s in the box, he might wish he had chosen differently. He could end up discovering that the box contains chocolates he likes more than the cookies, or that it contains chocolates he likes less than the cookies.

*Question:* Which of the following general reasons best explains Edward’s less-than-complete confidence?

**Correct answer:** You Will Learn More After Making Your Decision.

**Understanding Question 3 of 7.**

Sarah is considering moving to a distant city. If she moves she’ll have exciting new opportunities, but she will miss her family and current friends. If she stays put she’ll have the comfort of familiarity, but she will find the lack of new experiences a bit stifling. Sarah thinks there are good arguments both for prioritizing new opportunities, and good arguments for prioritizing comfort and familiarity. If she prioritizes new experiences, moving is clearly better; if she prioritizes comfort and familiarity, staying is clearly better. She considers both ways of setting her priorities equally reasonable. Ultimately she decides to stay. She reports overall confidence less than 100 in her choice because she recognizes that either one of these perspectives – prioritizing new experiences, and prioritizing comfort and familiarity – would be completely appropriate for her to adopt.

*Question:* Which of the following general reasons best explains Sarah’s less-than-complete confidence?

**Correct answer:** Multiple Valid Views About What You Want.

**Understanding Question 4 of 7.**

Steven is choosing between two different exercise routines: one that promotes cardio fitness through jogging and calisthenics, and one that promotes muscle development through weight lifting. Steven understands the benefits of strength and stamina. He asks himself, which is more important to him? He thinks there's a correct answer to that question, and that he could figure it out if he gave the issue enough attention. Instead, because he's pressed for time, he quickly chooses the cardio fitness routine. He reports overall confidence less than 100 in his choice because he's not sure the routine he chose will deliver the benefits he would consider most important if he gave the issue enough attention.

*Question:* Which of the following general reasons best explains Steven's less-than-complete confidence?

**Correct answer:** You Worry You Could Have Made a Better Decision Without Access to Additional Information.

**Understanding Question 5 of 7.**

Sabrina is deciding which of two snacks will be delivered to her office next week: a candy bar or an apple. She knows that she usually prefers apples, but occasionally she craves a candy bar. Her objectives are clear from the start: she would ideally like to have a candy bar if she ends up craving one and an apple if she doesn't. However, she has to choose in advance and can't ask for a different snack to be delivered depending on whether she ends up having a craving. She selects the apple because, given everything she knows when she has to place her order, including the fact that she usually ends up wanting an apple, she thinks an apple is clearly her best choice. However, she reports overall confidence less than 100 in her choice because she knows she might end up craving a candy bar next week, and if she does she'll be disappointed that she receives the apple instead.

*Question:* Which of the following general reasons best explains Sabrina's less-than-complete confidence?

**Correct answer:** You Will Learn More After Making Your Decision.

**Understanding Question 6 of 7.**

Shana chose between two restaurants for lunch: a pizza place and a sushi place. She would order a pepperoni pizza at the pizza place and a spicy tuna roll at the sushi place. She has eaten both many times and knows how well she enjoys them. She thinks pepperoni pizza tastes better but that sushi is healthier. When she weighs these two aspects of her experience, she realizes that, overall, she desires pizza and sushi just about the same. She chooses the pizza place thinking she desires it just a little bit more overall, but expresses less-than-complete confidence because it's such a close call.

*Question:* Which of the following general reasons best explains Shana's less-than-complete confidence?

**Correct answer:** It's a Toss-Up.

**Understanding Question 7 of 7.**

Jude is ordering a salad for lunch through a phone app. He must choose whether to have dressing mixed in or on the side. He intends to order the salad with dressing on the side, but he's in a rush and clicks through the app quickly. After submitting his order, he thinks he may have accidentally clicked the box for dressing mixed in. He knows that if he had taken his time, he would be sure that he had placed his order correctly, in which case he wouldn't be worried. Because he rushed when he didn't have to, he expresses overall confidence in his choice less than 100.

*Question:* Which of the following general reasons best explains Jude's less-than-complete confidence?

**Correct answer:** You Worry You Could Have Made a Better Decision Without Access to Additional Information.

**A.2. Study-Context Classification Test.** The same Risky Unmixed task is shown in all nine questions. The experiment randomly selects this task from the five tasks in which Option B has an expected value \$1 greater than Option A. Each question uses a different name, randomly assigned without replacement from Alex, Avery, Casey, Jamie, Jordan, Morgan, Quinn, Riley, and Taylor. In the templates below, *[Name]*, *[subject pronoun]*, *[object pronoun]*, and *[possessive adjective]* denote these participant-specific substitutions. The scenario is shown first; the question and answer options appear after the subject clicks Next. The first stage ends with the instruction, "When you are ready, click Next to answer the question."

**Study-Context Example 1 of 9.**

This example involves [Name], who selected Option B and expressed less-than-complete confidence because the riskiness of Option B made the alternatives hard to understand. [Name] wondered: What payment would [subject pronoun] receive if [subject pronoun] chose Option B and a purple ticket were drawn? What payment would [subject pronoun] receive if [subject pronoun] chose Option A and a blue ticket were drawn? In each case, how many other possible payments were there? What were the chances [subject pronoun] would end up with each possible payment? [Name] thought [subject pronoun] had not fully appreciated all of the possible contingencies.

*Question:* Which general reason best fits this example?

**Correct answer:** You Worry You Could Have Made a Better Decision Without Access to Additional Information.

**Study-Context Example 2 of 9.**

This example involves [Name], who selected Option B and expressed less-than-complete confidence because [subject pronoun] was not sure how anxious [subject pronoun] would feel about the outcome after making [possessive adjective] choice, or how bad [subject pronoun] would feel if a gold ticket were drawn and [subject pronoun] received nothing.

*Question:* Which general reason best fits this example?

**Correct answer:** You Will Learn More After Making Your Decision.

**Study-Context Example 3 of 9.**

This example involves [Name], who selected Option B and expressed less-than-complete confidence because [subject pronoun] worries that [subject pronoun] didn't think hard enough about what was important to [object pronoun] when choosing between risky and safe options before making [possessive adjective] decision. To be clear, [subject pronoun] is sure [subject pronoun] could have determined what makes the options better or worse for [object pronoun] if [subject pronoun] had given the question enough thought. However, because [subject pronoun] made the decision more casually, [subject pronoun] wonders, did [subject pronoun] place the appropriate weight on average payoffs? On the variability of payoffs? On the best-case and worst-case outcomes? [Name] believes [subject pronoun] could have settled this issue by thinking hard enough before making [possessive adjective] decision, but [subject pronoun] did not put in the necessary time and effort.

*Question:* Which general reason best fits this example?

**Correct answer:** You Worry You Could Have Made a Better Decision Without Access to Additional Information.

**Study-Context Example 4 of 9.**

This example involves [Name], who selected Option B and expressed less-than-complete confidence because [subject pronoun] recognized that different judgments about the tradeoff between average payment and variability were legitimate for [object pronoun]. [Name] thinks it would be legitimate for [object pronoun] to place greater importance on the average payment, in which case Option B is clearly better. But [subject pronoun] also thinks it would be legitimate for [object pronoun] to place greater importance on low variability, in which case Option A is clearly better. There are many ways to evaluate the tradeoff between average payment and variability that make sense to [Name].

*Question:* Which general reason best fits this example?

**Correct answer:** Multiple Valid Views About What You Want.

**Study-Context Example 5 of 9.**

This example involves [Name], who selected Option B but expressed less-than-complete confidence because the riskiness of Option B made [object pronoun] unsure which option would ultimately turn out to have been the better choice in retrospect. [Name] did not think [subject pronoun] could have improved [possessive pronoun] choice using only the information [subject pronoun] had when deciding, because [subject pronoun] did not know whether a purple ticket or a gold ticket would be drawn from the bag. However, [subject pronoun] knew that if the drawn ticket turned out to be purple, [subject pronoun] would realize that option B would have been better in retrospect, and if the drawn ticket turned out to be gold, [subject pronoun] would realize that option A would have been better in retrospect.

*Question:* Which general reason best fits this example?

**Correct answer:** You Will Learn More After Making Your Decision.

#### Study-Context Example 6 of 9.

This example involves [Name], who selected Option B and expressed less-than-complete confidence because [subject pronoun] was not sure that the simplifying rule [subject pronoun] used had led to the correct choice. [Name] wanted to choose the option that would give the largest amount on average, but did not want to take the time to calculate the average. Instead, [subject pronoun] picked the option with the largest potential payoff.

*Question:* Which general reason best fits this example?

**Correct answer:** You Worry You Could Have Made a Better Decision Without Access to Additional Information.

#### Study-Context Example 7 of 9.

This example involves [Name], who selected Option B and expressed less-than-complete confidence because [subject pronoun] thought the two options were roughly equally attractive. [Name] was pretty certain about the overall attractiveness of each option. Although [subject pronoun] selected Option B, [subject pronoun] thought there was a chance [subject pronoun] could be a little happier with Option A.

*Question:* Which general reason best fits this example?

**Correct answer:** It's a Toss-Up.

#### Study-Context Example 8 of 9.

This example involves [Name], who selected Option B and is sure it was the best choice [subject pronoun] could have made, but expressed less-than-complete confidence because the riskiness of Option B made [object pronoun] unsure how much money [subject pronoun] would receive. [Name] did not think [subject pronoun] could have improved her choice using only the information she had when deciding, because [subject pronoun] did not know whether a purple ticket or a gold ticket would be drawn from the bag.

*Question:* Which general reason best fits this example?

**Correct answer:** You Will Learn More After Making Your Decision.

#### Study-Context Example 9 of 9.

This example involves [Name], who selected Option B and expresses less-than-complete confidence because [subject pronoun] is worried that [subject pronoun] may have made a mistake when calculating the average payments from the ticket colors, chances, and payments shown in the choice task. [Name] was convinced, and remains convinced, that the best option is the one that pays the largest amount on average. [Name] was aware, and remains aware, that all the information needed to calculate the average payment was available when [subject pronoun] made [possessive adjective] choice. [Name]'s concern is that [subject pronoun] may have added or multiplied numbers incorrectly.

*Question:* Which general reason best fits this example?

**Correct answer:** You Worry You Could Have Made a Better Decision Without Access to Additional Information.