Discussion Paper No. 1236

ISSN (Print) 0473-453X
ISSN (Online) 2435-0982

EXPERIMENTAL EVALUATION OF RANDOM INCENTIVE SYSTEM UNDER AMBIGUITY

Tomohito Aoyama
Nobuyuki Hanaki

March 2024

The Institute of Social and Economic Research
Osaka University
6-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan

Experimental Evaluation of Random Incentive System under Ambiguity*

Tomohito Aoyama[†] Nobuyuki Hanaki[‡]

March 27, 2024

Abstract

The random incentive system (RIS) is a standard incentive scheme used to elicit preferences in economic experiments. However, it has been speculated that RIS may not be incentive compatible when participants are concerned about ambiguity, i.e., that the choices observed under RIS do not reflect the underlying preferences. To examine the performance of RIS under ambiguity, we conducted three experiments online and in a laboratory. The results of the experiments suggest that RIS is incentive compatible. We argue that presenting choice situations in isolation may improve the incentive compatibility of RIS. We also argue that using RIS, together with an experimental guideline called Prince, may reduce the observed ambiguity aversion.

KEYWORDS: Random incentive system, Incentive compatibility, Ambiguity, Prince

JEL CLASSIFICATION: C91, D81

*We gratefully acknowledge financial support from Joint Usage/Research Center at Institute of Social and Economic Research (ISER), Osaka University, and Japan Society for the Promotion of Science (18K19954, 20H05631, 20K22088, 23H00055). Experiments reported in this paper were approved by IRB of ISER, Osaka University.

[†]Waseda University, Japan. E-mail: ta5492@gmail.com

[‡]Institute of Social and Economic Research, Osaka University, Japan, and University of Limassol, Cyprus. Email: nobuyuki.hanaki@iser.osaka-u.ac.jp

1. Introduction

Since Ellsberg (1961), many studies have attempted to identify individuals' ambiguity attitudes by experimentally implementing the Ellsberg paradox (Trautmann and van de Kuilen, 2015). The Ellsberg paradox consists of several tasks in which participants choose between ambiguous and risky bets. To incentivize each task, rewards are often paid according to the choice in a randomly selected task. This incentive scheme, called the random incentive system (RIS), is commonly used with the expectation that it is incentive compatible, i.e., that the observed choices reflect the underlying preferences.[1] However, it has been speculated that RIS may not be incentive compatible if participants have nontrivial attitudes toward ambiguity (Oechssler and Roomets 2014, Bade 2015, Kuzmics 2017). If this speculation is true, it is difficult to identify ambiguity attitudes. The present study experimentally evaluates the incentive compatibility of RIS under ambiguity by comparing choices observed using and not-using RIS. In particular, we test monotonicity, a preference condition considered by Azrieli et al. (2018), which is sufficient for RIS to be incentive compatible. The theoretical background of the experiment is provided in Appendix A.

The first experiment (the main experiment, hereafter) was conducted online and consisted of two sessions that were two weeks apart. In each session, participants were presented with two boxes. Box K contained nine green balls and 11 yellow balls, while Box U contained 20 balls that were either red or black, with no information provided on the number of each color. There were two choice situations in which participants were asked to choose between the bets on green and red and between bets on green and black. We manipulated the incentive schemes. In the control treatment, called ONE, only one choice situation was incentivized. In the other treatment, called TWO, both choice situations were incentivized with RIS. Some participants were assigned to ONE in both sessions, and others were assigned to different treatments among the sessions. We compare the choices in the incentivized choice situations both between and within participants. In the between-participant comparison, ONE serves as the benchmark. If there is a difference in choices between ONE and TWO, it can be concluded that RIS is not incentive compatible. In the within-participant comparison, participants assigned to ONE in both sessions serve as the benchmark. As they experienced the same treatment twice, any difference in their choices between the sessions are interpreted as errors. In contrast, the choices of participants who experienced

[1]Other incentive schemes have been used in the experimental economic literature. For example, participants are sometimes paid for all choice situations. Azrieli et al. (2018) show that this mechanism requires a stronger preference condition to be incentive compatible than RIS.

both treatments might change their choices between the sessions because of the difference in the incentive schemes. If the propensity to change choices between sessions differs between these groups, it can be concluded that RIS is not incentive compatible. In both comparisons, we found no difference. Therefore, the results of the experiment support the incentive compatibility of RIS.

Baillon et al. (2022a) conducted an experiment with the same objective as ours. They found that the use of RIS drastically reduces ambiguity-averse behavior, which indicates that RIS is not incentive compatible. That is, their experimental result and ours are divergent. While their experimental design and ours differ in many respects, we focus on two differences. The first is the difference in the designs of the control treatments. While the control treatment used by these authors differed from their other treatments in the choice procedures, our control treatment ONE is similar to our TWO except for the incentive schemes. The second is that while they followed Prince, an experimental guideline developed by Johnson et al. (2021), we did not. To investigate the effect of these differences, we conducted additional experiments.

Our second experiment (the additional online experiment, hereafter) investigates the effect of the control treatment design and was conducted online. There were two treatments, SINGLE and ONE*. The treatment SINGLE was designed to resemble the control treatment of Baillon et al. (2022a). In SINGLE, participants first selected a choice situation to be faced with and then chose a bet presented in that choice situation. The treatment ONE* was a variant of ONE, where participants were presented with only one choice situation. We compared the proportions of participants who are identified as being ambiguity averse between SINGLE in this experiment and TWO in the main experiment, finding no significant difference. That is, we could not replicate the result of Baillon et al. (2022a) in this experiment. Additionally, we compared the observed choices between ONE* in this experiment and TWO in the main experiment, finding a significant difference. This result suggests that the framing of control treatment in the main experiment could influence the results.

The third experiment (the additional laboratory experiment, hereafter) aimed to conceptually replicate the experiment of Baillon et al. (2022a) and our main experiment in a laboratory, following Prince. This experiment included treatments ONE, TWO, and SINGLE, which were similar to those in the online experiments. We compared choices between ONE and TWO, and the proportions of participants identified as being ambiguity averse between SINGLE and TWO. We found no significant difference in either comparison, supporting the incentive compatibility of RIS. As a side result, we found that the proportion of participants in TWO who were identified as ambiguity averse is smaller in this experiment than in the main experiment. The use of RIS together with Prince may reduce the observed ambiguity aversion.

Overall, our experiments support the incentive compatibility of RIS, in contrast to Baillon et al. (2022a). This may be because we asked participants to answer each question on an individual page or sheet, while Baillon et al. (2022a) used one answer sheet. Presenting choice situations in isolation may retrieve the incentive compatibility of RIS.

1.1. Literature

Oechssler and Roomets (2014), Bade (2015), and Kuzmics (2017) speculated that RIS is not incentive compatible under ambiguity because it provides an opportunity to hedge ambiguity through randomness. Additionally, Baillon et al. (2022b) predicted that the incentive compatibility of RIS may be restored if the randomization occurs before decisions. Baillon et al. (2022a) conducted experiments to evaluate RIS under ambiguity, and their evidence suggests that it is not incentive compatible. The present paper adds to this literature by conducting experiments with different designs from that of Baillon et al. (2022a).

The incentive compatibility of RIS has also been questioned under risk. Holt (1986) and Karni and Safra (1987) predicted the failure of RIS under nonexpected utility preferences. Later, several experiments tested this prediction, and most authors have found evidence consistent with it (Cubitt et al. 1998, Cox et al. 2014, Harrison and Swarthout 2014, Cox et al. 2015, Brown and Healy 2018, and Freeman et al. 2019, Chew et al. 2022).

1.2. Organization

The rest of the paper is organized as follows. Section 2 describes the main experiment, while Sections 3 and 4 detail additional experiments conducted online and in a laboratory, respectively. Section 5 provides a general discussion, and Section 6 concludes the paper.

2. Main Experiment

The experiment was conducted online in two sessions, two weeks apart, in December 2021 using Qualtrics (`www.qualtrics.com`). Invitation emails were sent to students in the subject pool of Institute of Social and Economic Research (ISER), Osaka University, managed by ORSEE (Greiner, 2015). Among them, 413 participants completed Session 1, of which 312 also completed Session 2. Participants were instructed to complete the task individually by clicking on the link they received via email on the same day. On average, they took 156 sec. and 155 sec. to complete Sessions 1 and

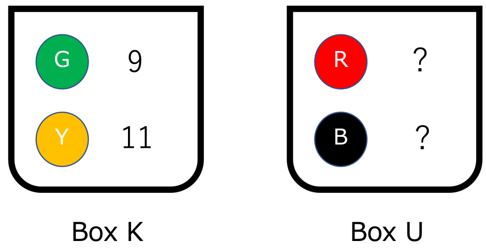


Fig. 1 The two boxes presented to participants

2, respectively. On average, they earned 329 JPY and 326 JPY in Sessions 1 and 2, respectively, including a participation fee of 100 JPY each. The reward was an Amazon gift card via email.[2]

2.1. Tasks

Participants were presented with two boxes, Box K and Box U, as depicted in Figure 1.[3] Box K contained nine green balls and 11 yellow balls.[4] Box U contained 20 balls that were either red or black, with the numbers of each color determined by students who participated in another experiment. Although participants were informed that the color composition was determined in this way, they were not informed of the exact number of each color.

Participants were presented with the following three bets.

Bet G: If a ball taken out of Box K is green, 500 JPY is paid out,

Bet R: If a ball taken out of Box U is red, 500 JPY is paid out,

Bet B: If a ball taken out of Box U is black, 500 JPY is paid out.

Then, they were asked to answer the following two questions:

Question R: Which do you prefer, Bet G or Bet R?

Question B: Which do you prefer, Bet G or Bet B?

These questions were displayed on individual pages.

Participants received instructions about the two questions before making their choices, so they could view the entire experiment as a single decision problem.

[2]This experiment was preregistered on October 30, 2021: `https://aspredicted.org/DFY_K5Y`.

[3]The two boxes were referred to as Box A and Box B in the experiment, respectively.

[4]We used nine green balls instead of 10 for tie-breaking, which prevents ambiguity-neutral participants from choosing Bet G in both questions. This technique was used by Oechssler and Roomets (2015).

	Session 1	Session 2
Group A	ONE	TWO
Group B	TWO	ONE
Group C	ONE	ONE

Table 1 Groups of participants in the main experiment

2.2. Incentives

At the beginning of the experiment, either Question R or Question B was selected as the target of rewards, or the real choice situation (RCS).[5] The method of selecting the RCS varied across treatments. After participants made choices, one ball was drawn from each of Box K and Box U, which were simulated on the computer. Then, participants received payments based on the outcome of the bet they chose in the RCS.

2.3. Treatments

There were two treatments, ONE and TWO, where the selection procedures of the RCS differed.

In ONE, the RCS was the question displayed later. For instance, if Question R appeared first and then Question B appeared, the RCS would be Question B. Participants were informed of this selection method. The order of the questions was randomized across participants, resulting in Question R being the RCS for half of the participants and Question B for the other half. We informed participants about this randomization to prevent them from suspecting that we presented a bet with a lower probability of winning in the second question.

In TWO, the RCS was selected following RIS. Participants were informed that the computer randomly chose each question as the RCS with a probability of one half before answering the questions. The RCS was revealed after participants answered both questions.

2.4. Sessions

Participants were not informed beforehand that there would be two sessions. We recruited back those who participated in Session 1 for Session 2 and connected the data using their ID number.

Participants were randomly assigned to one of three groups that differed in the treatments in the two sessions, as summarized in Table 1. Group A was assigned to ONE in Session 1 and to TWO

[5]In all the experiments, we selected the RCS under RIS at the beginning, which might enhance the incentive compatibility of RIS, as argued by Baillon et al. (2022b).

in Session 2, while Group B was assigned to TWO in Session 1 and ONE in Session 2. Group C was assigned to ONE in both sessions.

The numbers of red and black balls in the two sessions were determined by different students. In Session 2, we informed participants of this fact to prevent them from making their decision based on the outcome of Session 1. While the order of Question R and Question B was randomized among participants, it remained fixed across the two sessions.

2.5. Main Variable

The variable of interest is the answer to the question displayed later. In ONE, participants were expected to report their preferences truthfully in response to that question because it was known to be the RCS. In TWO, participants might manipulate answers to hedge ambiguity through the randomness of RIS. If participants in these treatments answer the second question differently, we conclude that the use of RIS distorts their choices.

Both between and within-participant comparisons are conducted. For the between-participant comparison, we compare the proportions of participants who chose Bet G in the second question between ONE and TWO.

For the within-participant comparison, Group C serves as a benchmark. Participants in Group C were assigned to ONE in both sessions. If they answered the second question differently between the sessions, the difference is interpreted as being caused by errors. However, participants in Groups A and B were assigned to different treatments in the two sessions, so their choices may differ because of the difference in the incentive schemes, in addition to errors. If the proportion of participants in Groups A and B who answer differently in the two sessions is greater than that of Group C, we can conclude that the difference in incentive schemes caused participants to change their answers.

As noted in the Introduction, these comparisons are intended to test monotonicity, a preference condition considered by Azrieli et al. (2018), which is sufficient for RIS to be incentive compatible.

2.6. Results

As participants were invited to Session 2 after completing Session 1, there was attrition in the sample. As reported in Appendix B, there was no observed effect of the treatment as well as the outcome in Session 1 on the rate of attrition. Therefore, we pool the data from the three groups when possible. In Session 1, 257 participants were involved in ONE and 156 in TWO. In Session 2, 191 participants were involved in ONE and 121 in TWO.

The chi-squared test was used unless otherwise noted. Fisher's exact test was also used, but the results are not reported as they are similar to those obtained from the chi-squared test. A significance level of 5% was adopted for all tests.

We first report the result of the between-participant comparison. In Session 1, 64.6% and 68.6% of participants chose Bet G in the second question in ONE and TWO, respectively. This difference is not statistically significant ($p = 0.47$).[6] A similar analysis using the data from Session 2 also shows no significant difference ($p = 0.91$).[7]

Result 1 (Between Participant Comparison)**.** The use of RIS did not have a significant effect on choices.

We next proceed to the within-participant comparison. Of participants in Groups A and B, 31.3% changed their answers, while 30.6% of those in Group C did. These proportions were not significantly different ($p = 1$).[8]

Result 2 (Within Participant Comparison)**.** The assignment to different treatments between sessions did not have a significant effect on their propensity to change their answers.

These results suggest that RIS is incentive compatible.

2.7. Discussion: Comparison with Baillon et al. (2022a)

Prior to conducting the main experiment, we became aware of Baillon et al. (2022a), whose experiment tests the validity of RIS under ambiguity. Their main finding is that the use of RIS drastically decreases the observed ambiguity aversion, suggesting that RIS is not incentive compatible.

There were many differences between our main experiment and that of Baillon et al. (2022a). Among them, we focused on the following two differences. First, the designs of the control treatments differed. In our experiment, treatment ONE was similar to treatment TWO except for the

[6]This analysis uses the data of all participants, including those who did not proceed to Session 2. Therefore, the sample sizes for ONE and TWO are 257 and 156, respectively.

[7]The proportions of participants who chose Bet G were 62.3% (119 out of 191) and 63.6% (77 out of 121) in ONE and TWO, respectively.

[8]This analysis uses the data of participants who completed Session 2. The sample size for Groups A and B is 240 and that for Group C is 75.

incentive schemes. However, in the control treatment Single of Baillon et al. (2022a), participants first selected one of two questions to answer (i.e., color to bet) and then chose between risky and ambiguous bets in response to the selected question. This choice procedure was different from that in treatments where RIS was used. Second, the formats of the experiments differed. While our main experiment was conducted online and computerized, theirs was conducted in a laboratory using pen and paper and physical devices. In particular, they followed the experimental guideline Prince proposed by Johnson et al. (2021). Prince calls for several principles to be followed, including the use of a tangible format for implementing RIS, which can potentially improve participants' clear understanding of the experiments. This guideline could also be a cause of the divergent results.

To examine the impact of these differences, we conducted an online experiment and a laboratory experiment, which are detailed in Sections 3 and 4, respectively.

3. Additional Online Experiment

The second experiment aimed to investigate the impact of using a benchmark treatment similar to Baillon et al. (2022a) and the effect of the way control treatment is framed in our main experiment. The tasks were similar to those in the main experiment, but the incentive schemes and framings differed.

The experiment was conducted online in January 2022 using Qualtrics. Invitation emails were sent to students in the same subject pool as in the main experiment.[9] Of the invited students, 206 completed the experiment. Participants were asked to complete the task individually by clicking on the individualized link they received via email on the same day. On average, participants took 124 sec. to complete the experiment. On average, they earned 360 JPY, including 100 JPY as a participation fee. Participants received their reward in the form of an Amazon gift card by email.[10]

3.1. Treatments

There were two treatments, SINGLE and ONE*, with differing incentive schemes.

SINGLE was designed to resemble the control treatment "Single" of Baillon et al. (2022a). Participants in SINGLE first selected either Question R or Question B as the RCS. They were then shown their selected question and chose a bet.

[9] Participants in the main experiment were excluded from the invitation.

[10] This experiment was preregistered on January 12, 2022: `https://aspredicted.org/3J1_Y7T`.

ONE* was a variation of ONE in the main experiment, differing only in the framing. In ONE*, participants were provided with only one question, which was the RCS. As in ONE, we informed participants in ONE* that the RCS is Question R for half of them and Question B for the other half.

3.2. Main Variable

The variable of interest is the answer to the displayed question.

The main finding of Baillon et al. (2022a) is that participants in the treatment with RIS were less likely to exhibit strictly ambiguity-averse (SAA) behavior compared with those in "Single." To re-examine this finding, we compare the observed ambiguity attitudes using data of SINGLE in this experiment and TWO in the main experiment. Participants in SINGLE are classified as SAA if they chose Bet G in the RCS. Participants in TWO are classified as SAA if they chose Bet G in response to both questions. The proportions of SAA participants in these treatments are compared.

Additionally, we examine the effect of the way control treatment is framed in our main experiment. For this purpose, we compare the proportion of participants who chose Bet G in ONE* in this experiment with the proportion of participants who chose Bet G in the second question in TWO of the main experiment.

3.3. Results

The number of participants in ONE* and SINGLE were 103 each. We first compare SINGLE and TWO. In SINGLE, the proportion of SAA participants was 58.3%. In TWO, it was 58.3% in Session 1 and 53.7% in Session 2. There was no significant difference in the proportions of SAA participants between SINGLE and TWO ($p = 1$ for Session 1 and $p = 0.59$ for Session 2).

Result 3 (Re-examination of Baillon et al. 2022a)**.** The use of RIS did not have a significant effect on the observed ambiguity aversion.

Thus, the main finding of Baillon et al. (2022a) was not replicated in this experiment.

We next compare ONE* and TWO. As reported in Section 2, 68.6% of participants in TWO of Session 1 answered Bet G is preferred in response to the second question. However, 52.4% of participants in ONE* chose Bet G. This difference is significant ($p = 0.013$).

Result 4 (Comparison of ONE* and TWO)**.** Participants were less likely to choose Bet G in ONE* compared with TWO.

That is, using ONE* instead of ONE allows us to conclude that the use of RIS enhances SAA behavior, which is an opposite effect to that of Baillon et al. (2022a). This result suggests that the framing of control treatment in the main experiment could influence the results.

4. Additional Laboratory Experiment

The third experiment was intended to examine the overall effect of experimenting in a laboratory and following Prince, the guideline followed in the experiment of Baillon et al. (2022a). The tasks were similar to those in experiments reported in Sections 2 and 3.

This experiment was conducted at the experimental laboratory of Osaka University in July 2022. A total of 126 participants (students of Osaka University, i.e., the same subject pool as the main and the additional online experiments) participated in the experiment in seven sessions, excluding randomly selected assistants.[11] The average session duration was 20 minutes. On average, participants earned 1234 JPY, including 1000 JPY as a participation fee.[12]

4.1. Treatments

There were three treatments: ONE, TWO, and SINGLE. While we manipulated the incentive schemes in a similar way to the online experiments, they were implemented with physical devices as follows.

In ONE, the RCS was Question R for half of participants and Question B for the rest. Participants first wrote answers on the answer sheets for both questions. The assistant then drew a ball from each box and announced its color. Participants were paid according to the outcome of their bets in the RCS. The flow of the payment stage was similar in the other two treatments.

In TWO, at the beginning, the assistant privately rolled a six-sided die, three sides of which were red and the rest black, to determine the RCS for all participants in the session. The assistant then put sheets of paper with the question related to the color of the die in sealed envelopes. Participants received envelopes and were told that the RCS was written inside. They then answered both questions. Participants then opened the envelopes and received rewards according to the outcome of their bets in the RCS.

In SINGLE, participants had to select one question they wanted to answer. They selected a

[11] Participants in the main experiment and the additional online experiment were excluded from the invitation.

[12] This experiment was preregistered on July 9, 2022: `https://aspredicted.org/3N4_MB8`

question to be answered and then answered the selected question. Participants were paid according to the outcome of the bet they chose in the selected choice situation.

4.2. Main Variable

The main variables are choices in incentivized questions and inferred ambiguity attitudes.

We compare the choices made in ONE and TWO. As mentioned above, the RCS in ONE was Question R for half of the participants, and Question B for the other half. We call the former group of participants ONE-R and the latter ONE-B. We compare the proportions of participants who chose Bet G over Bet R (resp. Bet B) between ONE-R (resp. ONE-B) and TWO.

We also compare the proportions of participants exhibiting SAA behavior between SINGLE and TWO. Participants in TWO are identified as being SAA if they chose Bet G in both questions, and participants in SINGLE are identified as being SAA if they chose Bet G in the selected choice situation.

4.3. Results

The number of participants in ONE-R and ONE-B were 20 and 21, respectively. Those in TWO and SINGLE were 41 and 44, respectively. We first compare the data from ONE and TWO. The proportion of participants who preferred Bet G to Bet R was 75% in ONE-R and 44% in TWO, respectively. The proportion of participants who preferred Bet G to Bet B was 33% in ONE-B and 51% in TWO, respectively. The p-values, corrected using the Bonferroni method, for these differences are $p = 0.088$ and $p = 0.57$, respectively. Therefore, the differences are not significant.

Result 5 (Between-Participant Comparison in Lab)**.** The use of RIS did not have a significant effect on choices.

The propensity to choose the risky bet (Bet G) over an ambiguous bet (Bet R or Bet B) in the RCS is significantly greater in ONE-R than in ONE-B ($p = 0.018$). This observation cannot be explained by a bias in participants' belief that Box U contained many black balls because this is not consistent with the data from TWO and SINGLE. In fact, participants in TWO did not have a different propensity to choose Bet G in the two questions ($p = 0.37$, McNemar test). Furthermore, the number of participants in SINGLE who chose Bet R and Bet B is the same. We could not find an explanation for the difference between ONE-R and ONE-B other than an error.

We next compare SINGLE and TWO. The proportion of SAA participants is 50% in SINGLE and 34.1% in TWO. This difference is not significant ($p = 0.21$). [13]

Result 6 (Reexamination of Baillon et al. (2022a) in Lab)**.** The use of RIS did not have a significant effect on the observed ambiguity aversion.

Thus, the main finding of Baillon et al. (2022a) was not replicated in this experiment.

Finally, we report a side result on the observed ambiguity aversion. In this experiment, the proportion of SAA participants in TWO, 34.1%, is significantly different from that in the main experiment, 58.3% ($p < 0.01$).

Result 7 (Comparison of Ambiguity Attitudes in Online and Lab Experiments)**.** The proportion of participants who exhibit SAA behavior under RIS is significantly lower in the lab than online.

5. General Discussion

Here, we discuss the implications of our experimental results.

5.1. Incentive Compatibility of Random Incentive System

In all experiments, we found no clear evidence that using RIS distorts choices. This is in contrast to the main finding of Baillon et al. (2022a) that the use of RIS reduces observations of SAA behavior. Several differences in the experimental designs might lead to these divergent results.

Perhaps the most important difference is how the questions were presented to participants. Whereas Baillon et al. (2022a) asked participants to write down answers to all questions on a single sheet of paper, we asked them to answer each question on individual pages or sheets. It is possible that presenting questions in isolation restores the validity of RIS. Consistently, in the context of choice under risk, Brown and Healy (2018) demonstrated that RIS distorts choices when an entire multiple price list is presented on a single page but not when each two-alternative choice is presented on an isolated page.

There are also two more subtle differences. First, the winning prize paid by Baillon et al. (2022a) was around 3.2 times larger than ours. Second, Baillon et al. (2022a) used a different tie-breaking method from ours. They reduced the winning prize of the risky bet, while we made

[13]The numbers of participants in SINGLE and TWO were determined so that the power of this test exceeds 0.95 under the effect size obtained from the data of Baillon et al. (2022a).

the winning probability of the risky bet slightly smaller than 50%. These differences might have also affected the results.

While the online experiments clearly support the incentive compatibility of RIS, the result of the lab experiment is less conclusive. As the lab experiment involved fewer participants than Baillon et al. (2022a), the possibility remains that the effect of RIS they found exists with an effect size smaller than in their experiments, perhaps because of the presentation of the questions in isolation. Further studies would be needed to clarify this point.

5.2. Observed Ambiguity Aversion

The proportion of SAA participants in TWO were 58.3% and 53.7% in Sessions 1 and 2 of the main experiment, respectively. However, the proportion in our laboratory experiment was much lower at 34.1%. Interestingly, Baillon et al. (2022a) reported that 28.7% of participants were SAA in a treatment similar to TWO, which is close to the proportion we found in the laboratory experiment.

Let us compare these proportions with those reported in existing studies. Oechssler and Roomets (2015) surveyed experimental studies focusing on the proportions of ambiguity averse participants, and reported a median of 58.5%. Among their survey subjects, eight studies used RIS, and the median obtained from these studies is 50.5%.[14] While the proportions of SAA participants in our main experiment are close to these medians, those in our laboratory experiment and of Baillon et al. (2022a) are much smaller.

This deviation from existing studies may be due to the guideline Prince, which Baillon et al. (2022a) and our lab experiment followed. Although more experimental evidence is needed for a better understanding, using RIS together with Prince may reduce the observed ambiguity aversion.

6. Conclusion

In this study, we investigated the incentive compatibility of RIS under ambiguity. We conducted online and laboratory experiments in which participants chose risky and ambiguous bets, under manipulation of the incentive schemes. We found no clear evidence that RIS distorts choice under ambiguity. This conclusion contrasts with that of Baillon et al. (2022a), who demonstrated that the use of RIS decreases observed ambiguity aversion. We identified the presentation of choice

[14]RIS was used by Akay et al. (2012), Butler et al. (2013), Cettolin and Riedl (2010), Charness et al. (2013), Chew et al. (2017), Dimmock et al. (2013), Keck et al. (2014), and Stahl (2014). Once we include studies that use variants of RIS (some participants are not actually rewarded or there are two or more RCSs), the median is 53.5%.

situations in isolation as a key difference that might lead to the contrasting result. As a secondary finding, it is suggested that RIS reduces the observed ambiguity aversion together with Prince.

References

Akay, Alpaslan, Peter Martinsson, Haileselassie Medhin, and Stefan T. Trautmann (2012) “Attitudes toward Uncertainty among the Poor: An Experiment in Rural Ethiopia,” *Theory and Decision*, 73 (3), 453–464.

Anscombe, F. J. and R. J. Aumann (1963) “A Definition of Subjective Probability,” *The Annals of Mathematical Statistics*, 34 (1), 199–205.

Azrieli, Yaron, Christopher P. Chambers, and Paul J. Healy (2018) “Incentives in Experiments: A Theoretical Analysis,” *Journal of Political Economy*, 126 (4), 1472–1503.

Bade, Sophie (2015) “Randomization Devices and the Elicitation of Ambiguity-Averse Preferences,” *Journal of Economic Theory*, 159, 221–235.

Baillon, Aurélien, Yoram Halevy, and Chen Li (2022a) “Randomize at Your Own Risk: On the Observability of Ambiguity Aversion,” *Econometrica*, 90 (3), 1085–1107.

——— (2022b) “Experimental Elicitation of Ambiguity Attitude Using the Random Incentive System,” *Experimental Economics*, 25 (3), 1002–1023.

Brown, Alexander L. and Paul J. Healy (2018) “Separated Decisions,” *European Economic Review*, 101, 20–34.

Butler, Jeffrey V, Luigi Guiso, and Tullio Jappelli (2013) “Manipulating Reliance on Intuition Reduces Risk and Ambiguity Aversion,” *Working Paper*.

Cettolin, Elena and Arno Riedl (2010) “Delegation in Decision Making under Uncertainty. Are Preferences Incomplete?” *Working Paper*.

Charness, Gary, Edi Karni, and Dan Levin (2013) “Ambiguity Attitudes and Social Interactions: An Experimental Investigation,” *Journal of Risk and Uncertainty*, 46 (1), 1–25.

Chew, Soo Hong, Bin Miao, Qiang Shen, and Songfa Zhong (2022) “Multiple-Switching Behavior in Choice-List Elicitation of Risk Preference,” *Journal of Economic Theory*, 204, 105510.

Chew, Soo Hong, Bin Miao, and Songfa Zhong (2017) "Partial Ambiguity," *Econometrica*, 85 (4), 1239–1260.

Cox, James C., Vjollca Sadiraj, and Ulrich Schmidt (2014) "Asymmetrically Dominated Choice Problems, the Isolation Hypothesis and Random Incentive Mechanisms," *PLoS ONE*, 9 (3), e90742.

——— (2015) "Paradoxes and Mechanisms for Choice under Risk," *Experimental Economics*, 18 (2), 215–250.

Cubitt, Robin P., Chris Starmer, and Robert Sugden (1998) "On the Validity of the Random Lottery Incentive System," *Experimental Economics*, 1 (2), 115–131.

Dimmock, Stephen G, Roy Kouwenberg, Olivia S Mitchell, and Kim Peijnenburg (2013) "Ambiguity Attitudes and Economic Behavior," *Working Paper*.

Ellsberg, Daniel (1961) "Risk, Ambiguity, and the Savage Axioms," *The Quarterly Journal of Economics*, 75 (4), 643–669.

Freeman, David J., Yoram Halevy, and Terri Kneeland (2019) "Eliciting Risk Preferences Using Choice Lists," *Quantitative Economics*, 10 (1), 217–237.

Greiner, Ben (2015) "Subject Pool Recruitment Procedures: Organizing Experiments with ORSEE," *Journal of the Economic Science Association*, 1 (1), 114–125.

Harrison, Glenn W. and J. Todd Swarthout (2014) "Experimental Payment Protocols and the Bipolar Behaviorist," *Theory and Decision*, 77 (3), 423–438.

Holt, Charles A (1986) "Preference Reversals and the Independence Axiom," *American Economic Review*, 76 (3), 508–515.

Johnson, Cathleen, Aurélien Baillon, Han Bleichrodt, Zhihua Li, Dennie Van Dolder, and Peter P. Wakker (2021) "Prince: An Improved Method for Measuring Incentivized Preferences," *Journal of Risk and Uncertainty*, 62 (1), 1–28.

Karni, Edi and Zvi Safra (1987) ""Preference Reversal" and the Observability of Preferences by Experimental Methods," *Econometrica*, 55 (3), 675–685.

Keck, Steffen, Enrico Diecidue, and David V. Budescu (2014) "Group Decisions under Ambiguity: Convergence to Neutrality," *Journal of Economic Behavior & Organization*, 103, 60–71.

Kuzmics, Christoph (2017) "Abraham Wald's Complete Class Theorem and Knightian Uncertainty," *Games and Economic Behavior*, 104, 666–673.

Oechssler, Jörg and Alex Roomets (2014) "Unintended Hedging in Ambiguity Experiments," *Economics Letters*, 122 (2), 243–246.

——— (2015) "A Test of Mechanical Ambiguity," *Journal of Economic Behavior & Organization*, 119, 153–162.

Raiffa, Howard (1961) "Risk, Ambiguity, and the Savage Axioms: Comment," *The Quarterly Journal of Economics*, 75 (4), 690–694.

Stahl, Dale O. (2014) "Heterogeneity of Ambiguity Preferences," *Review of Economics and Statistics*, 96 (4), 609–617.

Trautmann, Stefan T. and Gijs van de Kuilen (2015) "Ambiguity Attitudes," in Keren, Gideon and George Wu eds. *The Wiley Blackwell Handbook of Judgment and Decision Making*, 89–116, Chichester, UK: John Wiley & Sons, Ltd.

A. Theoretical Background and Predictions

Here, we provide a theoretical background and identification assumptions under which our between and within-participant comparisons are valid.

Our argument builds on analyses of RIS by Azrieli et al. (2018) and Brown and Healy (2018). The identification assumption for the between-participant comparison follows Brown and Healy (2018). As an extension, we provide an identification assumption for the within-participant comparison.

A.1. Incentive Compatibility of RIS and Monotonicity

We introduce the setup, following Azrieli et al. (2018).

Let R, B, and G denote Bet R, Bet B, and Bet G, which appear in the main experiment, respectively. Then, $L = \{R, B, G\}$ is the set of bets. A participant's preference for bets is represented by a complete and transitive binary relation $\succeq$ over L, with $\succ$ denoting the asymmetric part. The experimenter wants to learn $\succeq$.

In the experiment, the participant is presented with two choice situations $D_1 = \{R, G\}$ and $D_2 = \{B, G\}$, from which they are asked to report a preferred bet. The list of reports in these choice situations is represented by an element of $D = D_1 \times D_2$.

Under RIS, the RCS is uncertain from the perspective of the participant. Let $\Omega = \{1, 2\}$ be a state space that represents this uncertainty. The RCS is D_1 in state 1 and D_2 in state 2. By announcing $(c_1, c_2) \in D$ as the best-preferred bets, the participant forms a compound lottery $\langle c_1, c_2 \rangle \in L^\Omega$ that is given by

$$\langle c_1, c_2 \rangle(\omega) = \begin{cases} c_1 \text{ if } \omega = 1, \\ c_2 \text{ if } \omega = 2. \end{cases}$$

The participant can manipulate the announcement to form a desirable compound lottery.

To describe the participant's manipulation of the announcement, we need to consider extensions $\trianglerighteq$ of $\succeq$ over compound lotteries. An extension $\trianglerighteq$ is a complete and transitive preference relation over L^Ω, with $\triangleright$ denoting the asymmetric part of $\trianglerighteq$. For each preference $\succeq$, the experimenter has in mind a set of its extensions $\trianglerighteq$ that are possible, and such extensions are said to be admissible. The experimenter observes reports preferred by $\trianglerighteq$ instead of $\succeq$.

The experimenter wants the participant to report bets preferred by $\succeq$ truthfully. We say a report $c^* \in D$ is truthful for $\succeq$ if for each $i = 1, 2$ and $c'_i \in D_i$, $c^*_i \succeq c'_i$. RIS is incentive compatible if a report is generated by $\trianglerighteq$ precisely when it is truthful.

Definition 1. RIS is incentive compatible if for any preference $\succeq$, any admissible extension $\trianglerighteq$, any $c^* \in D$ truthful for $\succeq$, and any $c' \in D$, $\langle c^*_1, c^*_2 \rangle \trianglerighteq \langle c'_1, c'_2 \rangle$, with $\langle c^*_1, c^*_2 \rangle \triangleright \langle c'_1, c'_2 \rangle$ whenever c' is not truthful.

Azrieli et al. (2018) characterized RIS using the next condition on $\trianglerighteq$.

Definition 2. $\trianglerighteq$ is a monotone extension of $\succeq$ if for any $c, c' \in D$ such that $c_i \succeq c'_i$ for all i, we have $\langle c_1, c_2 \rangle \trianglerighteq \langle c'_1, c'_2 \rangle$, with $\langle c_1, c_2 \rangle \triangleright \langle c'_1, c'_2 \rangle$ whenever there is some i for which $c_i \succ c'_i$.

We also say the participant $(\succeq, \trianglerighteq)$ satisfies monotonicity if $\trianglerighteq$ is a monotonic extension of $\succeq$. Monotonicity is a mild condition meaning that the participant prefers to make a report that yields a better bet in all states.

Azrieli et al. (2018) showed that RIS is incentive compatible under the assumption that the participant satisfies monotonicity. In addition, they show that RIS is essentially the only incentive compatible mechanism when any extension that satisfies monotonicity is admissible. In other words, if the experimenter only assumes that the participant satisfies monotonicity, RIS is essentially the only incentive scheme that is expected to be incentive compatible.

A.2. Reversal of Order and Violation of Monotonicity

Monotonicity can be violated under ambiguity, as argued by Azrieli et al. (2018).

For an illustration, suppose that the participant is SAA, meaning $G \succeq R$ and $G \succeq B$, and both states realize with a probability of one-half. If the participant satisfies monotonicity, they would have a preference $\langle G, G \rangle \trianglerighteq \langle R, B \rangle$, which leads to the announcement of (G, G).

However, the participant may announce (R, B) instead of (G, G). By announcing (R, B), they would form a compound lottery $\langle R, B \rangle$, which yields R and B with the probability of one-half, as depicted in Fig. 2 (a). This compound lottery resolves in the following order: first the objective uncertainty is resolved, then the ambiguous subjective uncertainty is resolved. As Raiffa (1961) argued, the participant can identify the lottery as the one where the order of uncertainty resolution is reversed, as depicted in Fig. 2 (b). This indifference to the timing of uncertainty resolution is termed as reversal of order (Anscombe and Aumann, 1963). Reversal of order implies that the participant values the compound lottery $\langle R, B \rangle$ as highly as the one that wins with the probability

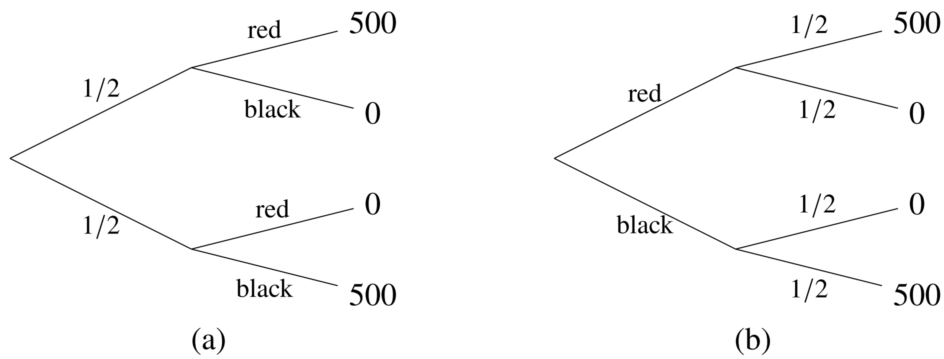


Fig. 2 Compound lotteries

of one half, irrespective of the color of the drawn ball. As the latter one stochastically dominates $\langle G, G\rangle$, the participant may have the preference $\langle R, B\rangle \rhd \langle G, G\rangle$, which violates monotonicity.

A.3. Predictions

Our experiments were intended to evaluate whether participants satisfy monotonicity by testing predictions that follow from it.

In both within- or between-participant comparisons, we assume that all participants satisfy the next condition, which is weaker than monotonicity.

Assumption 1. $(\succeq, \trianglerighteq)$ satisfies consistency if for any $c_i, c_j \in L$, $c_i \succeq c_j$ if and only if $\langle c_i, c_i\rangle \trianglerighteq \langle c_j, c_j\rangle$.

That is, the participant satisfies consistency if their preferences for degenerate compound lotteries are based on those for bets. This is a weak assumption without which the experimenter cannot expect to learn $\succeq$ by observing announcements induced by $\trianglerighteq$. Below, the participant is assumed to satisfy consistency, throughout.

Between-Participant Comparison

As argued by Brown and Healy (2018), the participant's preference $\succeq$ can vary according to the framing of the experiment. Thus, the performance of RIS should be evaluated under the control of the framing, and this is why we designed treatment ONE as close to treatment TWO as possible.

If the difference in incentive schemes also has a framing effect on $\succeq$, our comparisons do not work. Therefore, we need to exclude such an effect. For the between-participant comparison, we assume the next condition, which follows Brown and Healy (2018).

Assumption 2 (Mechanism Invariance)**.** $\succeq$ satisfies mechanism invariance if $\succeq$ does not differ between two experiments that are identical except for their incentive schemes.

We proceed to make a prediction tested by the between-participant comparison.

Prediction 1. Assume that all participants satisfy consistency and mechanism invariance. If all participants further satisfy monotonicity, then the proportions of participants who choose G in the second choice situation is the same in ONE and TWO.

Although the derivation of this prediction follows Brown and Healy (2018), we reproduce it for completeness below.

First, we argue that, under monotonicity, a participant's answer to the second question is the same whether the treatment is ONE or TWO. Imagine that a participant is in ONE, and suppose that their second question is Question B. Suppose further that they answer that question by choosing Bet G over Bet B.[15] Because Question B is the only incentivized question, this answer reveals that $\langle G, G \rangle \trianglerighteq \langle B, B \rangle$. Then, using the assumption of consistency, we infer that $G \succeq B$ in ONE. As the only difference between ONE and TWO is in the incentive schemes, from this preference and the assumption of mechanism invariance, we infer that the participant would also have preference $G \succeq B$ in TWO. Then, using the assumption of monotonicity, we infer that they would answer Question B by choosing Bet G over Bet B if they were in TWO.

As we randomly assign participants to ONE and TWO, the distributions of participants' preferences $\succeq$ among these treatments are considered to be the same. Thus, under monotonicity, we can predict that the proportion of participants who choose Bet G in the second choice situation is invariant between ONE and TWO.

Within-Participant Comparison

To analyze the within-participant comparison, we need to consider participants' preferences in the two sessions. We allow the preferences to differ between the sessions for unknown reasons and to vary with the framings. For $k = 1, 2$, let $(\succeq_k, \trianglerighteq_k)$ denote the pair of preferences for bets and compound lotteries in Session k.

Regarding the between-participant comparison, we need to assume that the preferences do not change with the incentive schemes in the two sessions.

[15] A similar argument applies to the case where the second question is Question R or the participant reports preference for an ambiguous bet.

Assumption 3 (Joint Mechanism Invariance)**.** $(\succeq_1, \succeq_2)$ satisfies joint mechanism invariance if the pair $(\succeq_1, \succeq_2)$ does not differ between two experiments that are identical except for their incentive schemes in the two sessions.

This is a stronger version of mechanism invariance.

We next make a prediction tested by the within-participant comparison.

Prediction 2. Assume that all participants satisfy consistency in both sessions and also satisfy joint mechanism invariance. If all participants further satisfy monotonicity, then the proportions of participants who answer the second question differently in the two sessions are the same for those who participate in ONE twice and those who participate in both treatments.

The derivation is similar to that of Prediction 1, as described below.

First, we argue that, under monotonicity, a participant's answers to the second question in the two sessions are independent of the treatments they are assigned. Imagine that the participant is assigned to ONE in both sessions and suppose that their second question is Question B. Suppose further that they answer that question by choosing Bet G over Bet B in both sessions. Similarly to the argument for Prediction 1, we infer that $G \succeq_1 B$ in ONE and $G \succeq_2 B$ in ONE. Consider how they would have answered the questions if the treatment in Session 1 and Session 2 were, say, ONE and TWO, respectively. As the only difference between ONE and TWO is in incentive schemes, using the assumption of joint mechanism invariance, we infer that their preferences in this case would be $G \succeq_1 B$ in ONE and $G \succeq_2 B$ in TWO. Then, using monotonicity, we infer that they would also answer Question B by choosing Bet G in both sessions in this case.

We randomly assign participants to Groups A and B, where treatments differ between the sessions, and Group C, where the treatment is ONE in both sessions. Thus, under monotonicity, we can predict that the proportions of participants who change answers to the second question between the sessions do not differ between the groups.

B. Sample Attrition in Main Experiment

Because we invited participants to Session 2 after they completed Session 1, there was some attrition of the sample. To examine the effect of the treatment and the outcome participants experienced in Session 1 on the rate of attrition, we conducted regression analyses to investigate the determinants of participation in Session 2. Table 2 shows the regression results.

The dependent variable is a dummy that takes the value of one if a participant proceeded to Session 2 and zero otherwise. The independent variables, all of which are dummies that take values of one or zero, are as follows.

- Green is one if participants answered Bet G is preferred to the second question of Session 1.
- Win is one if they won in Session 1.
- TWO is one if they were involved in TWO in Session 1.
- Order is one if the first question was Question B.

As can be seen from Table 2, none of these variables are significant even at the 10% significance level. Thus, we conclude there is no significant effect of the treatment or the outcome of Session 1 on participation in Session 2.

Furthermore, we tested several types of order effects. First, in each session, the order of questions did not affect answers to the second question. The p-values are $p = 0.46$ and $p = 1$ (resp. $p = 0.15$ and $p = 0.96$) for ONE and TWO in Session 1 (resp. in Session 2), respectively. Second, the order of treatments across the two sessions did not affect the answers to the second question in each treatment. The p-values are $p = 0.67$ and $p = 0.25$ for ONE and TWO, respectively. Overall, we found no significant order effect.

Table 2 Determinants of participation in Session 2

	Dependent variable:	
	Participation	
	OLS	*logistic*
	(1)	(2)
Green	0.012	0.067
	(0.045)	(0.243)
Win	0.070	0.387
	(0.044)	(0.240)
TWO	0.018	0.098
	(0.044)	(0.239)
Order	0.012	0.062
	(0.043)	(0.235)
Constant	0.702***	0.849***
	(0.049)	(0.257)
Observations	413	413
R^2	0.007	
Adjusted R^2	−0.002	
Log Likelihood		−228.187
Akaike Inf. Crit.		466.374
Residual Std. Error	0.431 (df = 408)	
F Statistic	0.769 (df = 4; 408)	
Note:		*p<0.1; **p<0.05; ***p<0.01

C. Instructions for Main Experiment

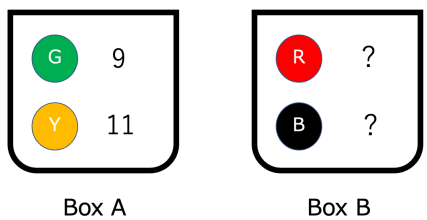
The following are the instructions translated from Japanese. Because the introduction is common to all treatments, we only show that for ONE.

C.1. Instruction for ONE

Introduction

Thank you for participating in this experiment. If you finish this experiment, we will pay you 100 yen as a participation fee. Depending on the results of the decisions you make during the experiment, you may be paid 500 yen in addition to the participation fee. Please note that we will not be able to pay you if your response time for the entire experiment exceeds 30 minutes.

Experiment description



Two boxes are simulated on a computer. Both boxes contain 20 balls. Box A contains nine green balls and 11 yellow balls, a total of 20. Box B contains red balls and black balls, a total of 20. The number of red and black balls in Box B was decided by a participant in another experiment that was conducted recently. We will not tell you the numbers of red and black balls.

Later, you will be asked to indicate your preferences for the following three lotteries.

- Lottery Green: If a ball is taken out of **Box A** and it is **green**, you are paid 500 yen, but if it is yellow, you are not paid anything.
- Lottery Red: If a ball is taken out of **Box B** and it is **red**, you are paid 500 yen, but if it is black, you are not paid anything.
- Lottery Black: If a ball is taken out of **Box B** and it is **black**, you are paid 500 yen, but if it is red, you are not paid anything.

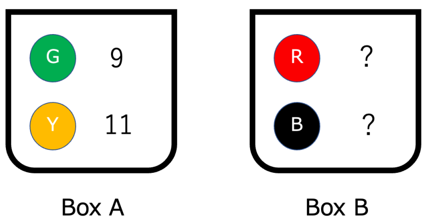
There are two questions. In both questions, we will present two of the three lotteries and ask you to choose one of them.

- Question 1: Which do you prefer, Lottery Green or Lottery Red?
- Question 2: Which do you prefer, Lottery Green or Lottery Black?

After these two questions, you will receive the lottery ticket you choose in Question 2. Then, a ball will be drawn from a box simulated on the computer, and if the lottery ticket is a winner, you will receive 500 yen in addition to the participation fee.

Half of the participants in this experiment will receive what they choose from Lottery Green and Lottery Red, and the other half will receive what they choose from Lottery Green and Lottery Black.

C.2. Instruction for TWO



Two boxes are simulated on a computer. Both boxes contain 20 balls. Box A contains nine green balls and 11 yellow balls, a total of 20. Box B contains red balls and black balls, a total of 20. The number of red and black balls in Box B was decided by a participant in another experiment that was conducted recently. We will not tell you the number of red and black balls.

Later, you will be asked to indicate your preferences for the following three lotteries.

- Lottery Green: If a ball is taken out of **Box A** and it is **green**, you are paid 500 yen, but if it is yellow, you are not paid anything.
- Lottery Red: If a ball is taken out of **Box B** and it is **red**, you are paid 500 yen, but if it is black, you are not paid anything.
- Lottery Black: If a ball is taken out of **Box B** and it is **black**, you are paid 500 yen, but if it is red, you are not paid anything.

There are two questions. In both questions, we will present two of the three lotteries and ask the participants to choose one of them.

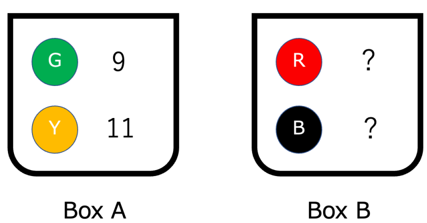
- Question 1: Which do you prefer, Lottery Green or Lottery Red?
- Question 2: Which do you prefer, Lottery Green or Lottery Black?

At the beginning of the experiment, the computer will randomly choose either Question 1 or Question 2 as the target of the reward. Both questions will be chosen with a probability of one-half. We will not tell you which question is the target of the reward until the reward payment stage. You will receive a lottery ticket for your choice on the question that is the target of the reward. Then, a ball will be drawn from the computer-simulated box, and if the lottery you receive is a winner, you will receive 500 yen in addition to your participation fee.

D. Instructions for Additional Online Experiment

The following are the instructions translated from Japanese. The instructions included an introduction that is similar to the introduction to the main experiment, which is omitted here.

D.1. Instruction for SINGLE



Two boxes are simulated on a computer. Both boxes contain 20 balls. Box A contains nine green balls and 11 yellow balls, a total of 20. Box B contains red balls and black balls, a total of 20. The number of red and black balls in Box B was decided by a participant in another experiment that was conducted recently. We will not tell you the number of red and black balls.

Later, you will be asked to indicate your preferences for the following three lotteries.

- Lottery Green: If a ball is taken out of **Box A** and it is **green**, you are paid 500 yen, but if it is yellow, you are not paid anything.

- Lottery Red: If a ball is taken out of **Box B** and it is **red**, you are paid 500 yen, but if it is black, you are not paid anything.

- Lottery Black: If a ball is taken out of **Box B** and it is **black**, you are paid 500 yen, but if it is red, you are not paid anything.

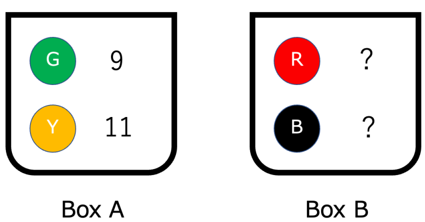
First, you will be asked which of the following questions you would like to answer.

- Question (Red): Which do you prefer, Lottery Green or Lottery Red?

- Question (Black): Which do you prefer, Lottery Green or Lottery Black?

If you respond that you want to answer Question (Red), Question (Red) will be displayed. The same applies to Question (Black). After you answer the displayed question, you will receive the lottery ticket that you choose. Then, a ball will be drawn from the computer-simulated box, and if the lottery you receive is a winner, you will receive 500 yen in addition to your participation fee.

D.2. Instruction for ONE*



Two boxes are simulated on a computer. Both boxes contain 20 balls. Box A contains nine green balls and 11 yellow balls, a total of 20. Box B contains red balls and black balls, a total of 20. The number of red and black balls in Box B was decided by a participant in another experiment that was conducted recently. We will not tell you the number of red and black balls.

Later, you will be asked to indicate your preferences for two of the following three lotteries.

- Lottery Green: If a ball is taken out of **Box A** and it is **green**, you are paid 500 yen, but if it is yellow, you are not paid anything.

- Lottery Red: If a ball is taken out of **Box B** and it is **red**, you are paid 500 yen, but if it is black, you are not paid anything.

- Lottery Black: If a ball is taken out of **Box B** and it is **black**, you are paid 500 yen, but if it is red, you are not paid anything.

There is only one question posed, as follows.

Question: Which lottery will you choose, Lottery Green or Lottery Black?

After you answer the question, you will receive the lottery ticket you choose. Then, a ball will be drawn from the computer-simulated box, and if the lottery you receive is a winner, you will receive 500 yen in addition to your participation fee.

Half of the participants in this experiment will be presented with Lottery Green and Lottery Red, and the other half will be presented with Lottery Green and Lottery Black.

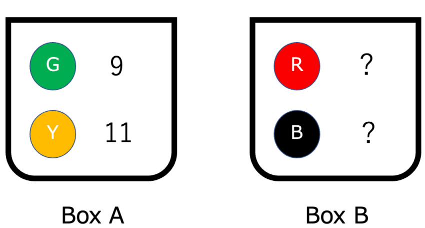
E. Instruction for Additional Laboratory Experiment

The following are the instructions translated from Japanese.

E.1. Instruction for ONE

Thank you for participating in this experiment. If you finish this experiment, we will pay you 1000 yen as a participation fee. Depending on the results of the decisions you make during the experiment, you may be paid 500 yen in addition to the participation fee.

There are two boxes. Box A contains nine green balls and 11 yellow balls, a total of 20. Box B contains red balls and black balls, a total of 20. The number of red and black balls in Box B was decided by a participant in another experiment that was conducted recently. We will not tell you the number of red and black balls.



Later, you will be asked to indicate your preferences for the following three lotteries.

- Lottery Green: If a ball is taken out of **Box A** and it is **green**, you are paid 500 yen, but if it is yellow, you are not paid anything.

- Lottery Red: If a ball is taken out of **Box B** and it is **red**, you are paid 500 yen, but if it is black, you are not paid anything.
- Lottery Black: If a ball is taken out of **Box B** and it is **black**, you are paid 500 yen, but if it is red, you are not paid anything.

There are two questions.

- Question (Red): Which do you prefer, Lottery Green or Lottery Red?
- Question (Black): Which do you prefer, Lottery Green or Lottery Black?

You will later receive one answer sheet for each question. Please write your answers to the questions on both of them.

Procedure for Reward Payment

Additional rewards will be paid based on the answers to Question (Red). After the experimenter collects the answer sheet, the assistant will draw one ball from Box A and one from Box B without looking inside the boxes. The assistant will announce the colors of both balls and record them on the whiteboard.

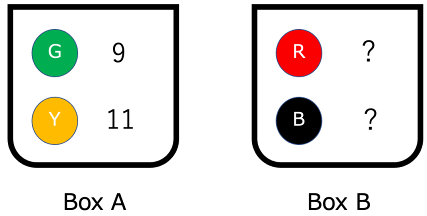
The rewards will be paid as follows.

- If you answer Question (Red) by choosing Lottery Green, you win if the assistant draws a green ball from Box A (45%).
- If you answer Question (Red) by choosing Lottery Red, you win if the assistant draws a red ball from Box B.

E.2. Instruction for TWO

Thank you for participating in this experiment. If you finish this experiment, we will pay you 1000 yen as a participation fee. Depending on the results of the decisions you make during the experiment, you may be paid 500 yen in addition to the participation fee.

There are two boxes. Box A contains nine green balls and 11 yellow balls, a total of 20. Box B contains red balls and black balls, a total of 20. The number of red and black balls in Box B was decided by the participants of another experiment that was conducted recently. We will not tell you the number of red and black balls.



Later, you will be asked to indicate your preferences for the following three lotteries.

- Lottery Green: If a ball is taken out of **Box A** and it is **green**, you are paid 500 yen, but if it is yellow, you are not paid anything.
- Lottery Red: If a ball is taken out of **Box B** and it is **red**, you are paid 500 yen, but if it is black, you are not paid anything.
- Lottery Black: If a ball is taken out of **Box B** and it is **black**, you are paid 500 yen, but if it is red, you are not paid anything.

There are two questions.

- Question (Red): Which do you prefer, Lottery Green or Lottery Red?
- Question (Black): Which do you prefer, Lottery Green or Lottery Black?

You will later receive one answer sheet for each question. Please write your answers to the questions on both of them.

Procedure for Reward Payment

Additional rewards will be paid based on the answers to the two questions. To decide which question is the target of the rewards, the assistant will roll a six-sided die once, with three sides painted red and three sides painted black. The question corresponding to the color of the roll will be the target of the rewards for all participants. You will be given an envelope containing a piece of paper with the reward question on it, and you will write your participant ID on it. Do not open the envelope until you are told to do so. Note that the question that determines the final payment amount is written on the paper inside the envelope and is determined before you answer the question. After you answer the questions, the experimenter will tell you to open the envelope. The assistant will then draw one ball from Box A and one ball from Box B without looking inside the boxes. The assistant will announce the color of both balls and record it on the whiteboard.

Based on the answers to the questions written on the paper inside the envelope, you will receive rewards as follows.
If Question (Red) is selected:

- In the case you choose Lottery Green, you win if the assistant draws the green ball from Box A (45%).
- In the case you choose Lottery Red, you win if the assistant draws a red ball from Box B.

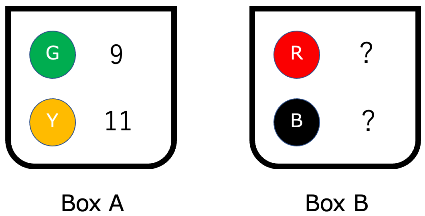
If Question (Black) is selected:

- In the case you choose Lottery Green, you win if the assistant draws a green ball from Box A (45%).
- In the case you choose Lottery Black, you win if the assistant draws a black ball from Box B.

E.3. Instruction for SINGLE

Thank you for participating in this experiment. If you finish this experiment, we will pay you 1000 yen as a participation fee. Depending on the results of the decisions you make during the experiment, you may be paid 500 yen in addition to the participation fee.

There are two boxes. Box A contains nine green balls and 11 yellow balls, a total of 20. Box B contains red balls and black balls, a total of 20. The number of red and black balls in Box B was decided by a participant in another experiment that was conducted recently. We will not tell you the number of red and black balls.



Later, you will be asked to indicate your preferences for the following three lotteries.

- Lottery Green: If a ball is taken out of **Box A** and it is **green**, you are paid 500 yen, but if it is yellow, you are not paid anything.

- Lottery Red: If a ball is taken out of **Box B** and it is **red**, you are paid 500 yen, but if it is black, you are not paid anything.
- Lottery Black: If a ball is taken out of **Box B** and it is **black**, you are paid 500 yen, but if it is red, you are not paid anything.

There are two questions.

- Question (Red): Which do you prefer, Lottery Green or Lottery Red?
- Question (Black): Which do you prefer, Lottery Green or Lottery Black?

You will later receive one answer sheet for each question. You will answer only one question. First, please choose which question you would like to answer. On either answer sheet, you will be asked if you wish to answer that question. On the answer sheet for the question you decide to answer, write your answer to that question.

Procedure for Reward Payment

Additional rewards will be paid based on the answers to a question. After the experimenter collects the answer sheet, the assistant will draw one ball from Box A and one from Box B without looking inside the boxes. The assistant will announce the colors of both balls and record them on the whiteboard.

Based on the answers to the questions you choose, you will receive rewards as follows.
If you choose Question (Red):

- In the case you answered Question (Red) by choosing Lottery Green, you win if the assistant draws a green ball from Box A (45%).
- In the case you answered Question (Red) by choosing Lottery Red, you win if the assistant draws a red ball from Box B.

If you choose Question (Black):

- In the case you answered Question (Black) by choosing Lottery Green, you win if the assistant draws a green ball from Box A (45%).
- In the case you answered Question (Black) by choosing Lottery Black, you win if the assistant draws a black ball from Box B.