The Virtual Bingo Blower: An open-source tool to generate ambiguity and risk in experiments

Online Appendix

Appendix A Instructions

When entering the experiment, subjects first saw a consent form. Then, they saw the instruction page.

A.1 Overall instructions

Welcome! In this experiment you will be asked to make various decisions. Depending on your decisions you may earn a sizeable amount of money that will be paid to you after the experiment. This will be composed of the £1 participation fee plus your 'earnings' from the outcome of your decisions.

These earnings will depend partly on the decisions you take during the experiment and partly on chance. So you will need to read these instructions carefully.

Outline of the experiment You will be asked 6 questions of the same type. In each question you will be presented with two options, Option A and Option B.

After you have answered all 6 questions, one of them will be selected at random, and the Option you chose in this question will be played for real.

Therefore, it is in your interest to answer each question as if that were the question to be played out.

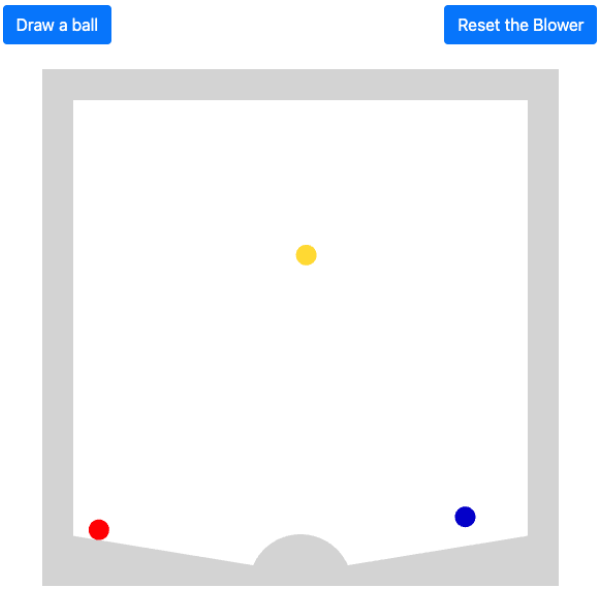


Figure A.1: The Bingo Blower as presented to participants, with the practice draw and reset buttons.

The Bingo Blower You will notice a virtual Bingo Blower below. You can see there are balls of three different colours: red, blue and yellow. It is a real physics simulation and not a pre-recorded video, meaning that we have no control on how the balls move.

To convince yourself that this Bingo Blower is simulated in real-time, on this page only you can manipulate the balls by clicking on one, dragging it and seeing how the other balls react. Later on you will not be able to control the balls to prevent you from affecting the outcome.

At the end of the experiment, in order to determine a colour, the Bingo Blower will be stopped and the ball closest to the center selected. Since the Bingo Blower is a real physics simulation, we cannot control the colour of the ball drawn.

You can use the buttons below to see how drawing a ball will work. After you click on 'Draw a ball', the balls keep tumbling for 3 seconds. Thereafter the balls are frozen and the ball closest to the center is chosen. If you want to try drawing a ball multiple times, click on 'Reset the Blower'.

[Figure A.1]

The ball drawn is: *[colour of the ball drawn]*

The questions In each question you will choose between two options, Option A and Option B:

- Option A (Bingo Blower option): you win £3 if the ball drawn from the Bingo

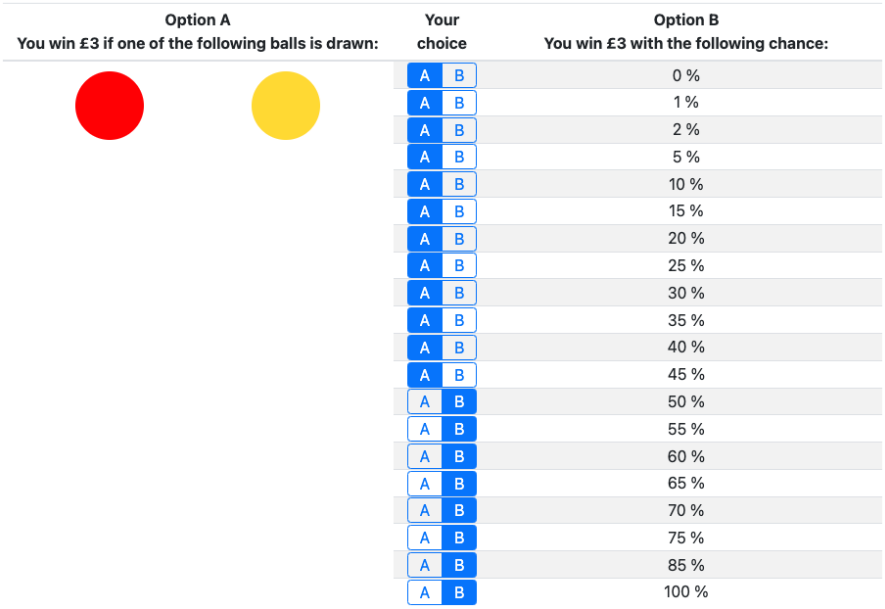


Figure A.2: Example choice list shown in the instructions.

Blower is of a certain colour, and nothing otherwise.

- Option B (given chance to win option): you have some chance to win £3.

You will be asked to state which one of these two options you prefer for various chances to win in Option B, from 0% to 100%. For example:

- If the chance to win in Option B is 100%, you will most likely prefer Option B because you are then guaranteed to win £3.
- If the chance to win in Option B is 0%, you will most likely prefer Option A because you might win something, as opposed to Option B where there is no chance to win at all.

You can think of the chance to win as the range of winning numbers in a bag containing numbers between 1 and 100. If the chance to win is 50%, it means you would win if a number between 1 and 50 is drawn from the bag, and you would lose if a number between 51 and 100 is drawn.

Example question The questions will look like this:

[Figure A.2]

In this example, Option A would lead to a gain of £3 if the ball drawn from the Bingo Blower is red or yellow. Option B is presented with ascending chances to win. For each line we ask you to choose between Option A and Option B.

If you prefer Option A when, for instance, Option B offers a 40% chance to win, then

you should also prefer Option A when Option B offers an even lower chance to win, say 25%. Similarly, if you prefer Option B with a 60% chance to win, then you should also prefer Option B with an even better chance to win, say 70%. Accordingly, in the figure above Option A is selected for every chance to win between 0 to 45%, and Option B, for every chance to win between 50% and 100%.

To save you the trouble of clicking for each line, you can click on Option A or Option B for one line and the computer will automatically fill your choices for the other lines. For example, in the figure above, button 'B' has been clicked only for the line corresponding to 50%.

You can change your choices as many times as you want. Once you are happy with your choices, click the 'Next' button at the bottom. You will not be allowed to go back to a question once you have clicked on the 'Next' button.

Payment After you have answered all 6 questions, one of them will be randomly selected by the computer. This question will be the question that counts for your payment. Then the computer will also select one of the lines:

- If, on this line, you had selected Option A, you will win £3 if the colour of the ball drawn from the Bingo Blower matches one of the colours displayed in Option A.
- If, instead, you had selected Option B, you will play that option. To do so the computer will randomly draw a number between 0 and 100. For example if Option B is 'You win £3 with a 74% chance' then you will win £3 if the number drawn by the computer is strictly smaller than 74, and you will win nothing if it is equal to or larger than 75.

A.2 Control questions

To test your understanding we have designed the following control questions.

Imagine one of the questions is as follows:

[Figure A.3]

Imagine further that this question is selected for payment at the end of the experiment.

Imagine that the line corresponding to a chance to win of 10% is selected. If your choices are as described in the example above, which option would you play at the end of the experiment? *[Single choice]*

- Option A
- Option B

What happens if a red ball is drawn? *[Single choice]*

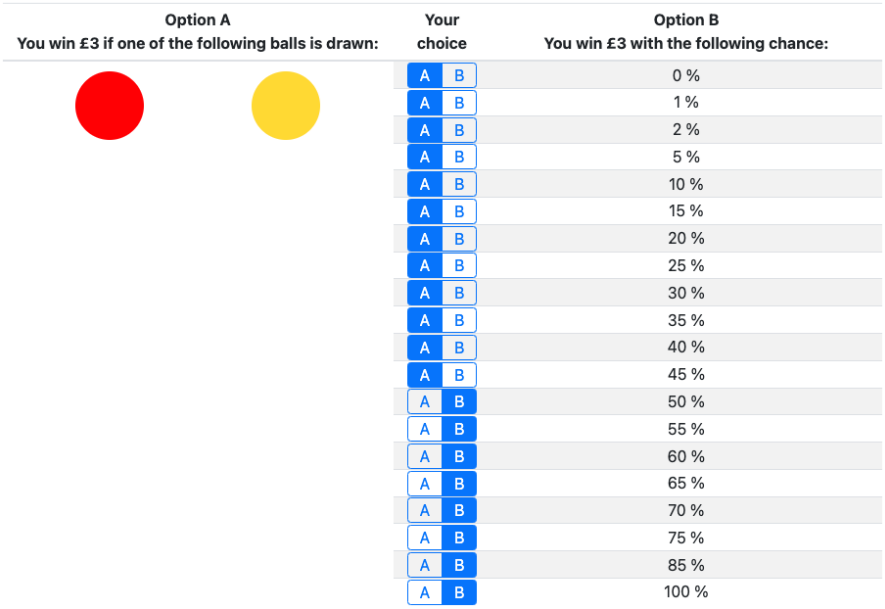


Figure A.3: Choice list used in the control questions.

- I gain £0
- I gain £3
- It is irrelevant: how much I gain depends on the random number selected by the computer to play Option B.

Imagine that the line corresponding to a chance to win of 60% is selected. If your choices are as described in the example above, which option would you play at the end of the experiment? *[Single choice]*

- Option A
- Option B

What happens if a red ball is drawn? *[Single choice]*

- I gain £0
- I gain £3
- It is irrelevant: how much I gain depends on the random number selected by the computer to play Option B.

A.3 Tasks

Subjects saw 6 questions, randomized independently for each subject.

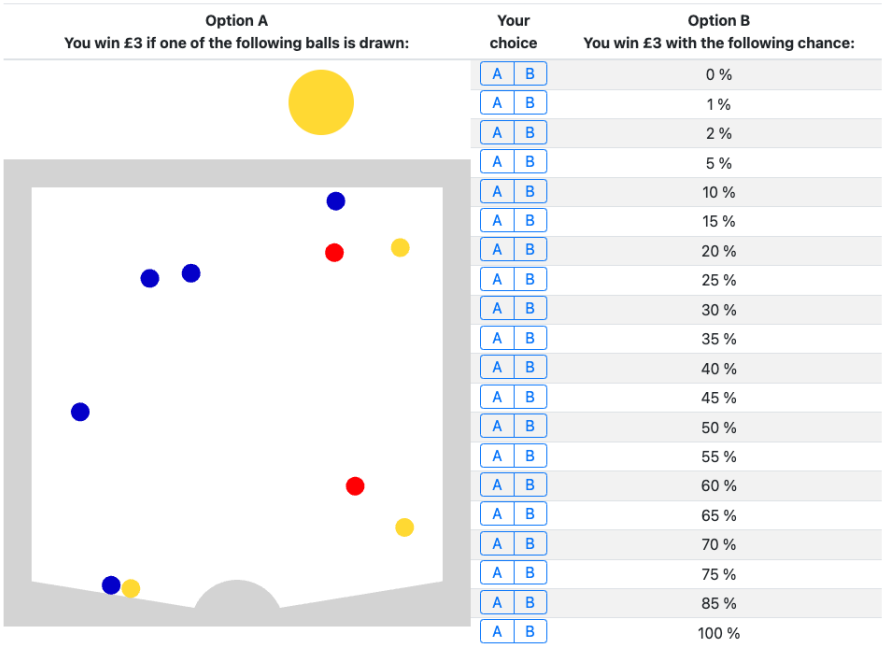


Figure A.4: Example of a choice task in VBB10, simple event.

A.3.1 VBB10 treatment

Figure A.4 gives an example of a task for a simple event, and Figure A.5, for a composite event.

A.3.2 VBB60 treatment

Tasks in the VBB60 treatment were identical to those in the VBB10 treatment except that the virtual bingo blower contained 60 balls.

Appendix B Summary statistics on the sample

Table B.1 describes the sample in terms of gender and age. As we can see there are subjects of varying ages present and slightly more females in the sample.

Appendix C Additional analyses

C.1 Accuracy of matching probabilities for composite events

In the main text, we report the accuracy of matching probabilities for the single events. Here, we report the corresponding analysis for the composite events.

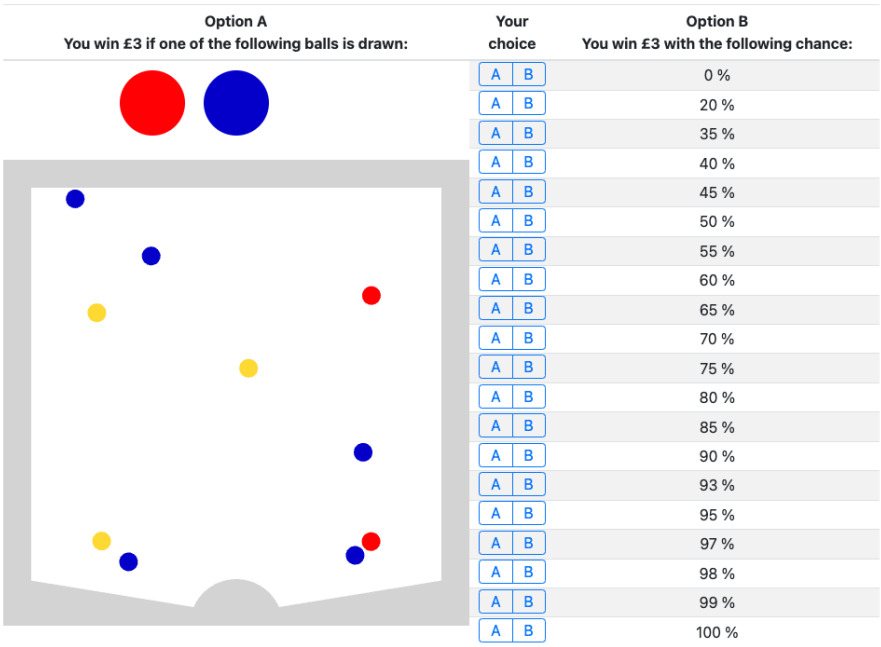


Figure A.5: Example of a choice task in VBB10, composite event.

Table B.1: Proportion of the different demographic categories by treatment

	VBB10	VBB60
Age:		
18–27	0.22	0.26
28–37	0.29	0.30
38–47	0.16	0.21
48–57	0.21	0.18
58+	0.12	0.06
Sex:		
Male	0.41	0.45
Female	0.59	0.55
Prefer not to say	0.00	0.00
Observations	162	155

As in the main text, we measure accuracy in three ways. First, the mean absolute error (MAE) between a subject's matching probabilities and the true probabilities of the composite events is 10.3 percentage points in VBB10 and 13.5 in VBB60 (Mann–Whitney test, $p < 0.001$). Second, the share of subjects who rank the three composite events in the right order is 50.0% in VBB10 and 19.4% in VBB60 (Fisher's exact test, $p < 0.001$). Ties count as a wrong ranking, and ranking at random would be right 16.7% of the time. Third, the share of subjects whose matching probabilities equal the true probabilities for all three composite events is 30.9% in VBB10 and 0.6% in VBB60 (Fisher's exact test, $p < 0.001$). For all six events, single and composite, it is 23.5% in VBB10 and 0% in VBB60 (Fisher's exact test, $p < 0.001$).

C.2 Distribution of the indices

The fact that only the a-index is affected by the treatments further confirms that these two indices measure different components of ambiguity attitudes. Baillon et al. (2018) already found that only the a-index is impacted by time pressure. We find a small negative correlation between the indices, significant at the 1% level (Pearson's ρ = -0.158, p = 0.005). Still, as can be seen in Figure C.1, which displays the a-index and the b-index of each subject as a scatter plot, the indices remain widely scattered without a strong pattern.

C.3 OLS regressions

Table C.1 reports the full OLS results. Models 1 and 2 are the preregistered specifications. Model 3 adds two controls that were not preregistered: the frame rate of the subject's device (FPS) and the time spent on the six tasks. The treatment effects barely change. The a-index is 0.088 higher in VBB60 than in VBB10 (one-sided $p = 0.060$), and the b-index still does not differ between treatments ($p = 0.632$).

C.4 Comments categorization

The categorization of comments was initially carried out using a pattern-based classification approach, implemented with a machine learning model (ChatGPT, OpenAI, used in September 2023) and Python.[1] This approach goes beyond simple keyword spotting and takes into account contextual clues and patterns in the text. The categories were initially suggested by the language model and were subsequently manually reviewed and adjusted

[1] See `https://bit.ly/46gujYR` for the full set of ChatGPT prompts used for the classification.

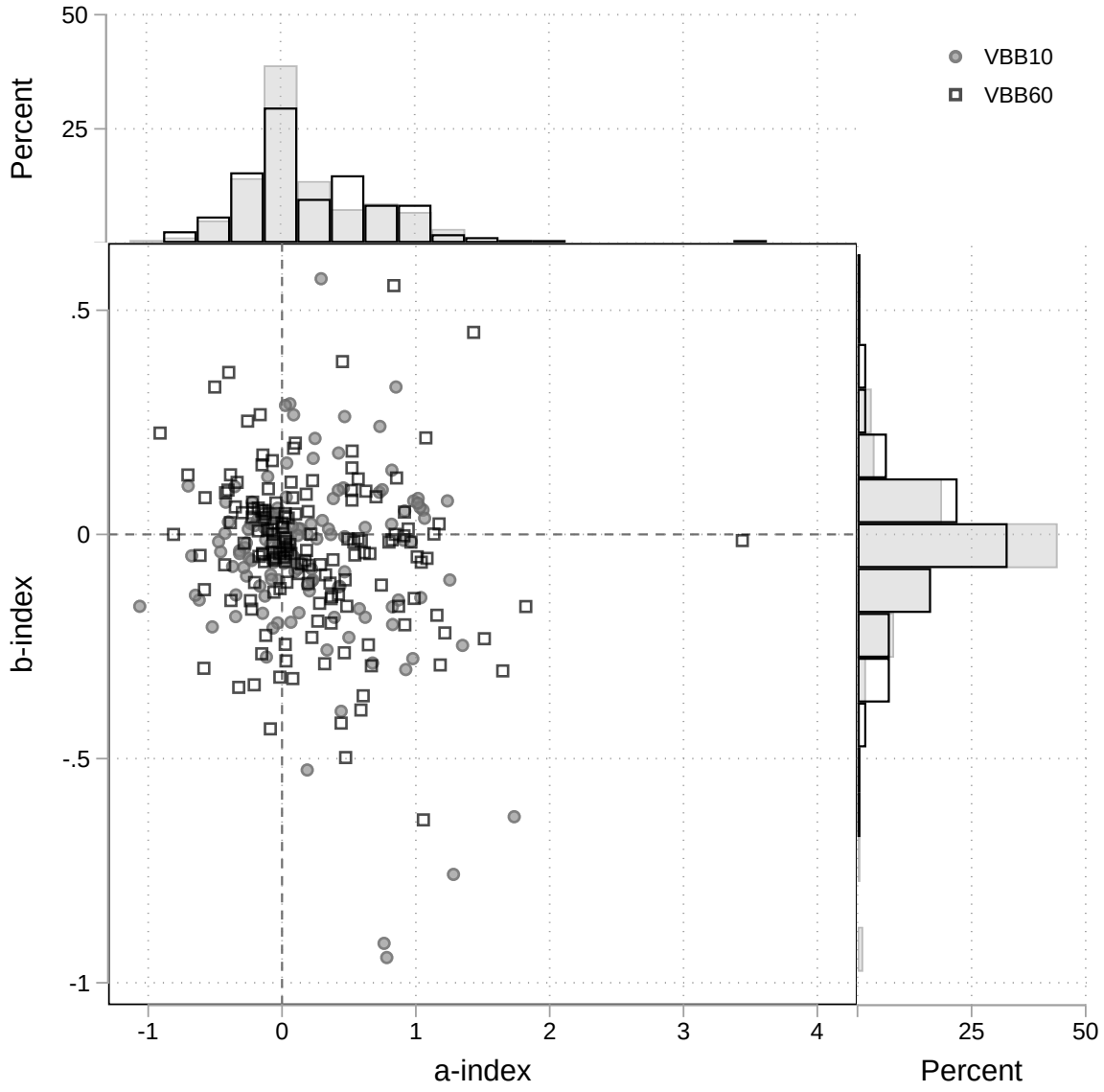


Figure C.1: a-index and b-index of each subject by treatment. Points in the scatter plot are slightly jittered.

Table C.1: Effect of treatments on a- and b-indices, OLS regressions.

	(1)		(2)		(3)	
	a-index	b-index	a-index	b-index	a-index	b-index
VBB60	0.073	0.009	0.094	0.007	0.088	0.010
	(0.058)	(0.020)	(0.058)	(0.020)	(0.057)	(0.021)
Age category (ref.: 18–27)						
28–37			0.169*	0.007	0.172*	0.009
			(0.084)	(0.028)	(0.084)	(0.028)
38–47			−0.038	0.015	−0.029	0.015
			(0.076)	(0.030)	(0.077)	(0.029)
48–57			0.115	0.011	0.131	0.013
			(0.090)	(0.032)	(0.092)	(0.033)
58+			0.259*	−0.027	0.270*	−0.025
			(0.106)	(0.054)	(0.107)	(0.055)
Female			0.013	0.004	0.002	0.008
			(0.058)	(0.020)	(0.058)	(0.022)
Time spent on tasks (minutes)					−0.033	0.005
					(0.021)	(0.010)
FPS					0.001	0.001
					(0.001)	(0.000)
Constant	0.176***	−0.047***	0.070	−0.053	0.078	−0.105
	(0.036)	(0.014)	(0.074)	(0.033)	(0.138)	(0.065)
Observations	317	317	317	317	317	317
R^2	0.005	0.001	0.041	0.005	0.049	0.009

Notes. Robust standard errors in parentheses.
Two-sided p-values. $^{\bullet}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. One-sided p-values for the a-index are reported in the text.

by the authors for accuracy. Based on this analysis, comments were assigned to one of the following categories:

- **Understanding:** Comments that mentioned confusion, difficulty in understanding, or explicitly stated incomprehension.
- **Engagement:** Comments that discussed interest, enjoyment, boredom, or difficulty in focusing on the experiment.
- **Technical Issues:** Comments that reported glitches, technical issues, or slow loading times.
- **Neutral/No comments:** Comments that explicitly stated no opinion or feedback.

We restrict attention to the 317 subjects used in the main analysis. A total of 63 comments were received in the experiment, out of which 17 were either neutral or had no comment. Table C.2 provides a summary of the categorized comments received from participants across different treatment groups. Each category is followed by the number of comments corresponding to it, broken down by treatment type (VBB10, VBB60). To offer a glimpse into the nature of the comments, representative examples for each category are also included in the table.

Appendix D Practical guidance for using the VBB

We can offer a number of practical solutions to mitigate any remaining fear of suspicion.

As we did for our experiments, subjects can be convinced that they face a real physics simulation by allowing them to move the balls in the VBB and see how the physics simulation reacts to their movement. This ability can be kept right until a ball needs to be drawn for payment. In addition, some elements of the VBB can be chosen by the subjects themselves before the purpose of the VBB is revealed, mirroring Oechssler and Roomets's (2015) construction of the Galton box. For example, subjects can be allowed to choose some of the parameters of the simulation, such as the speed of the wind or the weight of the balls.

Subjects can be asked to provide numbers that determine how long the balls continue moving in the VBB after clicking the button that starts the draw. The VBB can be further modified to allow subjects to click within the VBB to choose the location of the target from which the ball is drawn.

Table C.2: Categorization of comments.

Category	**VBB10**	**VBB60**	**Total**	**Example Comments**
Engagement	10	13	23	"I found it interesting! really enjoy these types of surveys.", "Wow that was exciting"
Neutral/No Comment	9	8	17	"no", "No further comments", "I have no comments"
Technical Issues	2	3	5	"It was difficult to count the number of each coloured ball as they were jumping all the time. There was also slightly delay for the program to response to clicks", "the balls move very fast so it was difficult to count how many of each there was in order to see the percentage", "Very interesting study. The page seemed to load rather slowly at times (even on pages without the balls present) but everything worked with patience. "
Understanding	10	8	18	" I had absolutely no idea what was going on despite reading the instructions several times. The description of Option B was completely baffling to me", "Although I did my best I apologize for struggling to fully understand the instructions. Thank you for giving me the opportunity to participate.", "Confusing design."
Total	**31**	**32**	**63**	

With some or all of these additional steps, it becomes even more transparent that the experimenter has no influence on the movement of the balls in the VBB and thus on the draw of a ball.

A final practical consideration concerns device performance. The VBB relies on the participant's own device to render the physics simulation. In the version of the VBB used in our experiment, the physics engine (`matter.js`) depended on the frame rate. Subjects with a lower frame rate therefore saw balls that moved more slowly and might become countable. This is why we excluded the few subjects whose frame rate fell below 50 FPS and why we control for the frame rate in our regressions. The physics engine has since been updated and no longer depends on the frame rate, as long as it stays above about 20 FPS. Device performance is therefore much less of a concern. Researchers who still want to check can record their subjects' frame rate, as we did in our experiments. The wiki of the VBB provides code snippets to do so.

References

Baillon, Aurélien, Zhenxing Huang, Asli Selim, and Peter P. Wakker (2018) "Measuring ambiguity attitudes for all (natural) events," *Econometrica*, 86 (5), 1839–1858, 10.3982/ECTA14370.

Oechssler, Jörg and Alex Roomets (2015) "A test of mechanical ambiguity," *Journal of Economic Behavior & Organization*, 119, 153–162, 10.1016/j.jebo.2015.08.004.