Appendix D. Experimental Instructions

of

"A Truth-Serum for Non-Bayesians: Correcting Proper Scoring Rules for Risk Attitudes,"

by

Theo Offerman, Joep Sonnemans, Gijs van de Kuilen, & Peter P. Wakker

The instructions have been translated from Dutch. The text between braces in the instructions for treatment ALL concerns the changes made in the instructions for treatment ALLnp.

Instructions treatment ONE

This experiment is about statements of which you do not know whether they are true or not. An example is the statement that snow did fall in Amsterdam in March 1861. You do not know for sure whether this statement is true or not. We will ask you to indicate how likely it is for you that such a statement is true, using probability judgments expressed in percentages. Perhaps you will, for example, attach a probability of 30% to the statement that it snowed in March 1861 in Amsterdam. We will then determine a score for you with the help of the added table on paper.

According to the table, for a probability judgment of 30% you get score 5100 if the statement is true (snow did fall in Amsterdam in March 1861). You get score 9100 if the statement is not true (snow did not fall in Amsterdam in March 1861). If you give a different probability judgment, you get different scores, as shown in the table. For example, if you give a probability judgment of 100%, your score is 10000 if the statement is true (snow did fall), and 0 if the statement is not true (snow did not fall). We now like to check whether the table with the scores is clear.

Practice questions using the table

Your answers were right. We will now explain some further features of the table. If you are certain that the statement is true, then it is best for you to give the maximum probability judgment of 100% because that gives the maximum score 10000 for a true statement. Every other answer then surely yields a lower score. If you are certain that the statement is not true, then it is similarly best to give the minimum probability judgment of 0%, because that gives the maximum score 10000 for a false statement. In many cases you do not know for certain whether a statement is true or not. We will now explain an important feature of the table on the basis of a thought experiment.

The properties of the table can be well illustrated with the help of repeated statements. Imagine, as a thought experiment, that you first have to give your probability judgment about a particular statement (for example, snow in Amsterdam in a particular year, say 1861). Imagine that you give judgment 30%, which means that you earn 5100 points in case of snow and 9100 points in case of no snow. Next however, various repetitions of that statement are being considered (snow in Amsterdam in March 1862, snow in Amsterdam in March 1863, …., snow in Amsterdam in March 1960), leading to a total of 100 of such statements. For all 100 statements (thus every year between 1861 and 1960) your score will be determined according to the table and your probability judgment (that is the same for every 100 statements). Your total score is then equal the sum of those 100 scores. For example, if it did snow in Amsterdam in March 35 times in those 100 years, and it did not snow 65 times, a probability judgment of 30% yields the following total score:

$$35 \times 5100 + 65 \times 9100 = 770000$$

We can also calculate this for other probability judgments, suppose that your probability judgment was 35%, then your total score was:

$$35 \times 5775 + 65 \times 8775 = 772500$$

On the next page we show that your total-score is optimal if your probability judgment is exactly equal to that percentage. Put differently, if for example 35 of the 100 (35%) statements are true, then it is best for you to choose probability judgment 35% because it will give you the highest total-score.

Now suppose that 35 of the 100 statements are true. We will determine what your total-score would have been at different judgments.

Table showing the total score for all possible probability judgments

It looks like judgment 35 is best. We conclude that if 35% of the statements are true, probability judgment 35 is optimal. Something similar holds for every percentage.

Conclusion: For every percentage of true statements your total-score is optimal if you choose your probability judgment to be equal to that percentage. Check this for another number by clicking on continue.

Subjects were required to check the conclusion for any other percentage

The experiment for non-repeated statements

The experiment we will perform concerns unique, and not repeated, statements. The various unique statements we consider are all different. For every single one of them you can give a different probability judgment.

There is a big difference between the real experiment and the thought-experiment with repetition. In the thought experiment there was an objective-optimal probability judgment, based on the percentage of true statements. In the real experiment, there are no repetitions and for every probability judgment you get only one score.

The thought experiment does give a guide for your probability judgment in the real experiment, with the percentage true statements as reference point. It is now based on your own subjective judgment however, and not on objective calculations. In the real experiment, there is no right or wrong answer. You purely choose what you like best.

In the experiment, you will encounter all different sorts of statements, more or less probable ones, and you can choose all probability judgments ranging from 0% till 100%. You can only choose whole percentages.

Payoff

This experiment consists of two parts. In both parts you will be asked to give probability judgments, 100 in part 1 and 20 in part 2. At the end of the experiment, one out of 120 statements considered during the experiment will be randomly (with equal probability) selected and on the basis of your score at this statement you will be paid out in euros, where 500 points is equal to 1 euro. Click on continue to read the instructions of the first part of the experiment.

Instructions part 1

In the graph below you see the price of a stock from January till June in a year in the past. We used real stock prices of the Amsterdam Exchange when we made the graphs. The graph is scaled in such a way that the price of the stock always stays between the upper and lower axis. The same holds for the other graphs you will see later in this experiment. We consider the following statement: on the 31st of December in that particular year, the price of the stock in the graph was in the purple area. We ask you to give a probability judgment about the truth of this statement without any further information about the stock or the year. You can only base this on the course of the graph in the first half of the year.

Figure showing an example of a graph of a stock price

Your score at this question depends on your probability judgment and whether the statement is true or not, according to the table.

Figure showing the same graph but with three different end prices at 31st of December

The input of your probability judgment takes place in two phases: first you type in an integer number between 0 and 100, next you will be shown a menu in which your choice is reproduced with the corresponding scores from the table. At that moment you can still alter your choice and choose any other integer between 0 and 100. You can do this by selecting the up or down arrow, or by clicking the mouse in the menu and scroll to another probability judgment. Next, when you click on OK your choice is final and you continue with the next statement. If you have any questions at this moment, raise your hand. The experimenter will come to you.

Instructions part 2

Part 1 of the experiment is now over. The second part of the experiment consists of 20 statements. Also in this part of the experiment you will be asked to give probability judgments. The difference is that it does not concern the prediction of stock prices now, but rolls with two 10-sided dice. On one of the dice are the values 00, 10, 20, 30, 40, 50, 60, 70, 80, 90 and on the other die are the values 1, 2, 3, 4, 5, 6, 7, 8, 9. Both dice will be rolled. The sum of the outcomes has the values 1-100 (we consider the roll 00-0 as if it is 100), where all values have the same probability.

Picture showing the two ten sided dice

An example of a statement is "the outcome is in the range 01-25." This statement is true when the outcome of the dice is indeed between 1 and 25 (including 25), and not true when the outcome is higher than 25. The input of your probability judgment again takes place in two phases: first you type in an integer number between 0 and 100, next you will be shown a menu in which your choice is replicated with the corresponding scores from the table. At that moment you can still alter your choice and choose any other integer number between 0 and 100. You can do this by selecting the up or down arrow, or by clicking the mouse in the menu and scroll to another probability judgment. Next, when you click on OK your choice is final and you continue with the next statement. Also in this part there is no right or wrong answer; you again choose what you want best. At the end of the experiment one statement will be selected and paid out. In case that this is a statement from part 2 of the experiment, you will be asked to roll the two ten sided dice once.

This is the end of part 2. Please raise your hand. The experimenter will come by so that it can be determined which round will be paid out.

Instructions treatment ALL [treatment ALLnp]

This experiment is about statements of which you do not know whether they are true or not. An example is the statement that snow did fall in Amsterdam in March 1861. You do not know for sure whether this statement is true or not. We will ask you to indicate how likely it is for you that such a statement is true, using probability judgments expressed in percentages [we will ask you to report a number]. Perhaps you will, for example, attach a probability of 30% to the statement that it snowed in March 1861 in Amsterdam [sentence deleted]. We will then determine a score for you with the help of the added table on paper. [We will then determine a score for you depending on the number you have reported with the help of the added table on paper.]

According to the table, for a probability judgment of 30% [number 30] you get score 5100 if the statement is true (snow did fall in Amsterdam in March 1861). You get score 9100 if the statement is not true (snow did not fall in Amsterdam in March 1861). If you give a different probability judgment [report a different number], you get different scores, as shown in the table. For example, if you give a probability judgment of 100% [report number 100], your score is 10000 if the statement is true (snow did fall), and 0 if the statement is not true (snow did not fall). We now like to check whether the table with the scores is clear.

Practice questions using the table

Your answers were right. We will now explain some further features of the table. If you are certain that the statement is true, then it is best for you to give the maximum probability judgment of 100% [maximum number 100] because that gives the maximum score 10000 for a true statement. Every other answer then surely yields a lower score. If you are certain that the statement is not true, then it is similarly best to give the minimum probability judgment of 0% [minimum number 0], because that gives the maximum score 10000 for a false statement. In many cases you do not know for certain whether a statement is true or not. We will now explain an important feature of the table on the basis of a thought experiment.

The properties of the table can be well illustrated with the help of repeated statements. Imagine, as a thought experiment, that you first have to give your probability judgment [number] about a particular statement (for example, snow in Amsterdam in a particular year, say 1861). Imagine that you give judgment 30% [report number 30], which means that you earn 5100 points in case of snow and 9100 points in case of no snow. Next however, various repetitions of that statement are being considered (snow in Amsterdam in March 1862, snow in Amsterdam in March 1863, …., snow in Amsterdam in March 1960), leading to a total of 100 of such statements. For all 100 statements (thus every year between 1861 and 1960) your score will be determined according to the table and your probability judgment [number] (that is the same for every 100 statements). Your total score is then equal the sum of those 100 scores. For example, if it did snow in Amsterdam in March 35 times in those 100 years, and it did not snow 65 times, a probability judgment of 30% [number 30] yields the following total score:

$$35 \text{ x } 5100 + 65 \text{ x } 9100 = 770000$$

We can also calculate this for other probability judgments [numbers], suppose that your probability judgment was 35% [number was 35], then your total score was:

$$35 \text{ x } 5775 + 65 \text{ x } 8775 = 772500$$

On the next page we show that your total-score is optimal if your probability judgment [number] is exactly equal to that percentage [the amount of times snow did fall]. Put differently, if for example 35 of the 100 statements (35%) [deleted] are true, then it is best for you to choose probability judgment 35% [report number 35] because it will give you the highest total-score.

Now suppose that 35 of the 100 statements are true. We will determine what your total-score would have been at different judgments [numbers].

Table showing the total score for all possible probability judgments

It looks like judgment [number] 35 is best. We conclude that if 35% of the statements [35 of the 100 statements] are true, probability judgment [number] 35 is optimal. Something similar holds for every percentage [number]. Conclusion: for every percentage of true statements your total-score is optimal if you choose your probability judgment [number] to be equal to that percentage [the amount of true statements]. Check this for another number by clicking on continue.

Subjects were required to check the conclusion for any other percentage [number]

The experiment for non-repeated statements

The experiment we will perform concerns unique, and not repeated, statements. The various unique statements we consider are all different. For every single one of them you can give a different probability judgment [number].

There is a big difference between the real experiment and the thought-experiment with repetition. In the thought experiment there was an objective-optimal probability judgment [number], based on the percentage [amount] of true statements. In the real experiment, there are no repetitions and for every probability judgment [number] you get only one score.

The thought experiment does give a guide for your probability judgment [number] in the real experiment, with the percentage [amount of] true statements as reference point. It is now based on your own subjective judgment however, and not on objective calculations. In the real experiment, there is no right or wrong answer. You purely choose what you like best.

In the experiment, you will encounter all different sorts of statements, more or less probable ones, and you can choose all probability judgments [numbers] ranging from 0% till 100% [0 till 100]. You can only choose whole percentages [numbers].

The experiment for non-repeated statements

The experiment we will perform concerns unique, and not repeated, statements. The various unique statements we consider are all different. For every single one of them you can give a different probability judgment [number].

There is a big difference between the real experiment and the thought-experiment with repetition. In the thought experiment there was an objective-optimal probability judgment

[number], based on the percentage of true statements. In the real experiment, there are no repetitions and for every probability judgment [number] you get only one score.

The thought experiment does give a guide for your probability judgment [numbers] in the real experiment, with the percentage [amount of] true statements as reference point. It is now based on your own subjective judgment however, and not on objective calculations. In the real experiment, there is no right or wrong answer. You purely choose what you like best.

In the experiment, you will encounter all different sorts of statements, more or less probable ones, and you can choose all probability judgments [numbers] ranging from 0% till 100% [0 till 100]. You can only choose whole percentages [numbers].

Payoff

This experiment consists of two parts. In both parts you will be asked to give probability judgments [report numbers], 100 in part 1 and 20 in part 2. Whether or not a statement was true will be revealed to you at the end of the experiment. Then, all 120 statements will be considered, and all scores will be determined. Your earnings in euro are equal to the sum of all scores divided by 60000. Click on continue to read the instructions of the first part of the experiment.

Instructions part 1

In the graph below you see the price of a stock from January till June in a year in the past. We used real stock prices of the Amsterdam Exchange when we made the graphs. The graph is scaled in such a way that the price of the stock always stays between the upper and lower axis. The same holds for the other graphs you will see later in this experiment. We consider the following statement: on the 31st of December in that particular year, the price of the stock in the graph was in the purple area. We ask you to give a probability judgment about the truth of this statement [number for this statement] without any further information about the stock or the year. You can only base this on the course of the graph in the first half of the year.

Figure showing an example of a graph of a stock price

Your score at this question depends on your probability judgment[the number you report] and whether the statement is true or not, according to the table.

Figure showing the same graph but with three different end prices at 31st of December

The input of your probability judgment [number] takes place in two phases: first you type in an integer number between 0 and 100, next you will be shown a menu in which your choice is reproduced with the corresponding scores from the table. At that moment you can still alter your choice and choose any other integer between 0 and 100. You can do this by selecting the up or down arrow, or by clicking the mouse in the menu and scroll to another probability judgment [number]. Next, when you click on OK your choice is final and you continue with the next statement. If you have any questions at this moment, raise your hand. The experimenter will come to you.

Instructions part 2

Part 1 of the experiment is now over. The second part of the experiment consists of 20 statements. Also in this part of the experiment you will be asked to give probability judgments [report numbers]. The difference is that it does not concern the prediction of stock prices now, but rolls with two 10-sided dice. On one of the dice are the values 00, 10, 20, 30, 40, 50, 60, 70, 80, 90 and on the other die are the values 1, 2, 3, 4, 5, 6, 7, 8, 9. Both dice will be rolled. The sum of the outcomes has the values 1-100 (we consider the roll 00-0 as if it is 100), where all values have the same probability.

Picture showing the two ten sided dice

An example of a statement is "the outcome is in the range 01-25." This statement is true when the outcome of the dice is indeed between 1 and 25 (including 25), and not true when the outcome is higher than 25. The input of your probability judgment [number] again takes place in two phases: first you type in an integer number between 0 and 100, next you will be shown a menu in which your choice is replicated with the corresponding scores from the table. At that moment you can still alter your choice and choose any other integer number between 0 and 100. You can do this by selecting the up or down arrow, or by clicking the mouse in the menu and scroll to another probability judgment. Next, when you click on OK your choice is final and you continue with the next statement. Also in this part there is no right or wrong answer; you again choose what you want best, and also in this part of the experiment, all scores will be summed and paid out. For convenience, you will therefore be asked to shake a box with 20 compartments each containing a pair of 10-sided dice at the end of the experiment. This box will then be opened, the result of each pair of dice will be inspected, and your earnings will be calculated on the basis of these results.

This is the end of part 2. The results of the lotteries will now be determined by shaking the box containing the pairs of dice. Please raise your hand so that the experimenter knows that you are ready.