# Communication of uncertainty in weather forecasts

Marimo, Pricilla and Kaplan, Todd R and Mylne, Ken and Sharpe, Martin

University of Exeter Business School

2012

**Communication of uncertainty in weather forecasts**

Pricilla Marimo[a1], Todd R. Kaplan[a], Ken Mylne[b] and Martin Sharpe[b]

[a] *University of Exeter, Business School, Streatham Court, Rennes Drive, Exeter, EX4 4PU,UK*

[b] *Met Office, FitzRoy Road Exeter, EX1 3PB, UK*

## Abstract

Experimental economics methods were used to assess public understanding of information in weather forecasts and test whether the participants were able to make better decisions using the probabilistic information presented in table or bar graph formats than if they are presented with a deterministic forecast. We asked undergraduate students from the University of Exeter to choose the most probable temperature outcome between a set of "lotteries" based on the temperature up to five days ahead. If they choose a true statement, participants were rewarded with a cash reward. Results indicate that on average participants provided with uncertainty information make better decisions than those without. Statistical analysis indicates a possible learning effect as the experiment progressed. Furthermore, participants who were shown the graph with uncertainty information took on average less response time compared to those who were shown a table with uncertainty information.

[1] Corresponding author e-mail: pm272@ex.ac.uk

## 1. Introduction

Providing probabilistic weather information to users has the potential to improve decision making, since weather is uncertain due to the chaotic nature of the atmosphere. Accordingly, the National Research Council (2006) state that a forecast is incomplete if uncertainty information is not included. The Met Office, through the use of ensemble forecasting and other techniques is capable of providing probabilistic estimates of weather forecasts. Studies that assess decision making when provided with probabilistic weather information have concluded that on average, participants who were given uncertainty information made significantly better decisions than those without (Roulston et al., 2006; Roulston and Kaplan, 2009). Still questions arise on whether or not the presentation format makes a difference in interpretation and understanding.

Different presentation formats/designs can be used to illustrate the same data in various fields. However, the way that information is presented and consequently how we interpret or process it has the potential to influence decision-making. Winett and Kagel (1984) note that although messages might contain the same information, the *format* and *modality* of presentation: visual, auditory or kinaesthetic (see, Fleming and Mills, 1992) and *context* in which the information is presented can have fairly different effects. Speier (2006) concluded that how information is presented and decision performance is moderated by the complexity of the task. Several studies have been done that focus on the impact of presentation format on decision making in the managerial or accounting field (Anderson and Mueller, 2005; Cardinaels, 2008; Ghani et al., 2009; So and Smith, 2003; Sullivan, 1988). In other studies, participants rate tabular reports as being less complex than graphical reports (Dickson et al., 1986; Lusk and Kersnick, 1979; Vessey and Galletta, 1991) whilst others prefer graphs to tables (Zmud, 1978). However when actual experiments are done, there are inconsistencies on which format is better. In some instances there are no differences in performance between subjects presented with tabular and graphical data (Benbasat and Dexter, 1985). In Remus (1984), both participants made costly decisions in the production scheduling problem despite whether or not they were provided with tabular or graphical displays. However when the erratic components of the decisions are reduced, the tabular aids outperformed the graphical aids. In other studies, graphical displays do better than tabular displays for example in an assessment of risk avoidance, participants who were shown graphical displays were willing to pay a higher price for improved toothpaste or set of four improved tires. These participants were also more willing to recommend others to buy improved tires compared to those who were shown numerical displays (Chua et al., 2006). Results from a study on communication of investment risk to consumers concluded that, presenting relative investment performance and probability of losing money in a bar chart instead of a table reduces customer's ability to correctly answer questions by between 50-75% (Driver et al., 2010).

The Met Office Public Weather Service (PWS) is constantly developing new products for disseminating weather information to users. After public consultation, they have a new format for presenting probabilistic forecast information for use on the Met office website. To make sure that the information is being communicated effectively to users in a way that will allow them to make better decisions it is desirable that methods of presenting the information are *objectively* evaluated. This study follows the same approach that was used by (Roulston and Kaplan, 2009). Their study tested the ability of subjects to understand the information in a fan chart format for expressing uncertainty in 5-day temperature forecasts. In our study, we test a table and bar graph format. We used experimental economics lab techniques to assess if participants make better decisions when provided with uncertainty information presented in a table and bar graph format than if they are presented with a deterministic forecast. The study will determine whether the method/format for communicating uncertainty information makes a difference on subject understanding of the forecast and test the speed at which subjects are able to learn with either method.

## 2. Experimental design/method

A total of 289 undergraduate students from various disciplines at the University of Exeter were recruited to participate in the experimental sessions. The sessions were computer based and took place in the Finance and Economics Experimental Laboratory (FEELE) at the University of Exeter. Experiments involved asking participants to choose between a set of "lotteries" based on the maximum temperature up to five days ahead. Participants were divided into three treatment groups: A, B and C. The 5-day temperature forecast information was presented as follows:

Group A: Table with a point forecast
Group B: Table with point forecast and uncertainty information
Group C: Bar graph with point forecast and uncertainty information

The uncertainty information that was provided for groups B and C showed the temperature range within the 90% confidence interval. The Group C format is currently and at the time of the experiment under trial on the Met Office website. Examples of the how the forecast information was presented for the three groups are shown in Figures 1 (a), (b) and (c).

(a)

| | Maximum Temperature (°C) | | | | |
|---|---|---|---|---|---|
| Most likely | 4 | 4 | 5 | 5 | 6 |
| | Sat 22 Jan | Sun 23 Jan | Mon 24 Jan | Tue 25 Jan | Wed 26 Jan |

(b)

| | Maximum Temperature (°C) | | | | |
|---|---|---|---|---|---|
| High range | 6 | 10 | 10 | 8 | 14 |
| Most likely | 4 | 4 | 5 | 5 | 6 |
| Low range | 3 | 2 | 2 | 2 | 0 |
| | Sat 22 Jan | Sun 23 Jan | Mon 24 Jan | Tue 25 Jan | Wed 26 Jan |

**Product description**

Table shows temperature ranges for a five day forecast. Temperatures fall within the indicated range roughly 9 out of 10 times with the most likely temperature in the middle. Presentation format is to help the Met Office develop and improve a product.

(c)

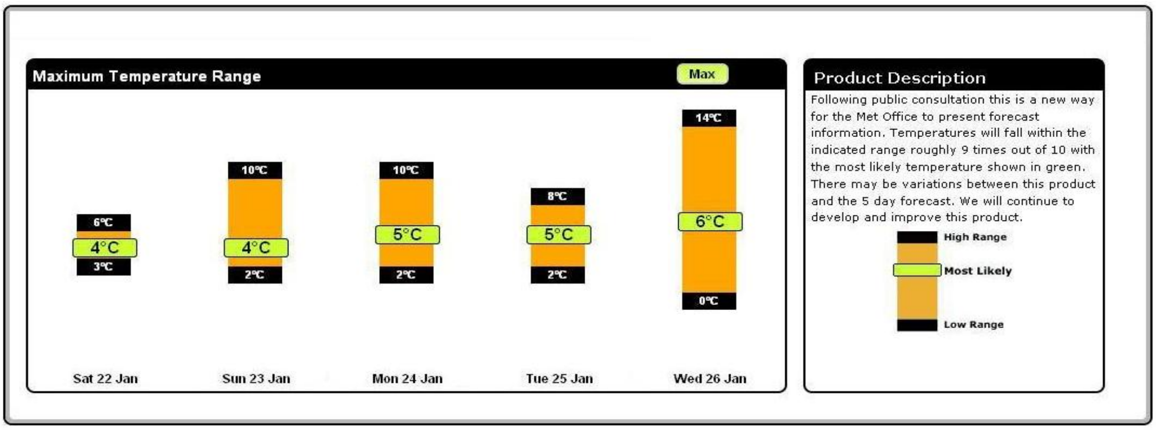



Figure 1. The participants were given the option to choose between two statements and receive £0.50.

Statement A – The maximum temperature on Saturday is above 5 deg. C OR

Statement B – The maximum temperature on Wednesday is above 8 deg. C

(a) The forecast presented to group A in question 4 of the experiment.

(b) The forecast presented to group B in question 4 of the experiment. The options were the same as those presented to group A

(c) The forecast presented to group C in question 4 of the experiment. The options were the same as those presented to group A

(a)

| | Maximum Temperature (°C) | | | | |
|---|---|---|---|---|---|
| Most likely | 5 | 6 | 5 | 1 | 3 |
| | Sat 1 Jan | Sun 2 Jan | Mon 3 Jan | Tue 4 Jan | Wed 5 Jan |

(b)

| | Maximum Temperature (°C) | | | | |
|---|---|---|---|---|---|
| High range | 6 | 10 | 14 | 3 | 11 |
| Most likely | 5 | 6 | 5 | 1 | 3 |
| Low range | 3 | 2 | 1 | -1 | -1 |
| | Sat 1 Jan | Sun 2 Jan | Mon 3 Jan | Tue 4 Jan | Wed 5 Jan |

**Product description**

Table shows temperature ranges for a five day forecast. Temperatures fall within the indicated range roughly 9 out of 10 times with the most likely temperature in the middle. Presentation format is to help the Met Office develop and improve a product.

(c)

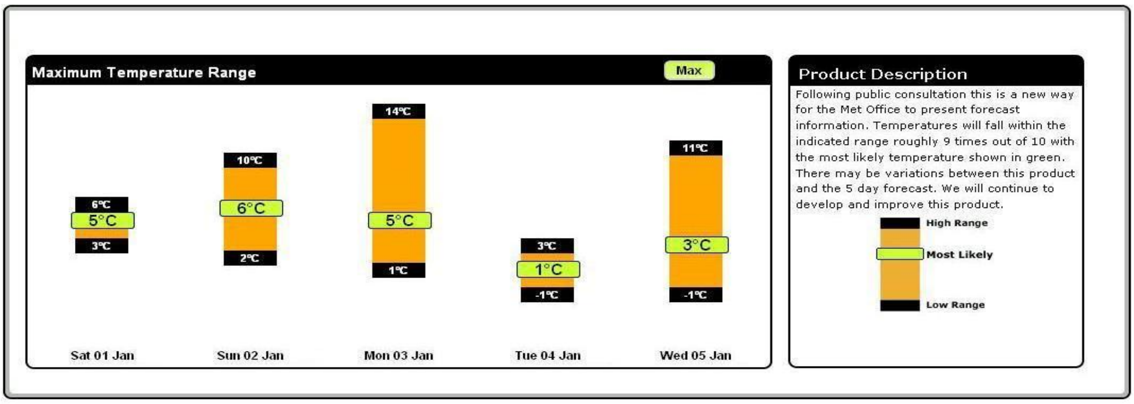



Figure 2. The participants were given the option to choose between two statements and receive £0.50.

Statement A – The maximum temperature Monday is above 6 deg. C OR

Statement B – The maximum temperature on Tuesday is above 0 deg. C

(a) The forecast presented to group A in question 1 of the experiment.

(b) The forecast presented to group B in question 1 of the experiment. The options were the same as those presented to group A

(c) The forecast presented to group C in question 1 of the experiment. The options were the same as those presented to group A

Instructions (shown in Appendix A) were provided on the computer screens and to assess whether or not participants could read the graphs and tables they had to answer three test questions at the beginning of the experiment. Two versions of test questions were used, one asked for the maximum temperature that is *'most likely'* whilst the other one asked for the maximum temperature that is *'expected'* (an example of a test question and number of participants shown each version shown in Appendix C). Ninety-three percent of the participants answered all the test questions correctly. Following Roulston and Kaplan (2009), all groups were presented with a sequence of 20 "lotteries" or rounds in which they had to choose the most probable temperature outcome based on the 5-day forecast. If the true statement was chosen, participants were rewarded with £0.50. The two criteria in each lottery had the following structure:

Statement A: The maximum temperature on Day D1 is above/below X deg. C

Statement B: The maximum temperature on Day D2 is above/below Y deg. C.

For each lottery, both the statements had the same preposition: both stated that the maximum temperature was "above" or both were "below" X/Y deg. C and none of them were "mixed" (i.e. one statement above and the other one below). This follows the same criterion that was used by (Roulston and Kaplan, 2009). After every round, the participants were shown a computer screen with the results from that particular round. This contained information on which statement the participant chose, which of the statement was true or false, the actual temperature for each day and their cumulative payoff. The same graphs were shown for all the participants in a particular group but in different orders. Four question orders were used and these were:

1$^{st}$ order: 1, 2,..., 20

2$^{nd}$ order: 20, 19,..., 1

3$^{rd}$ order: 11, 12,..., 20, 1, 2,..., 10

4$^{th}$ order: 10, 9,..., 1, 20, 19,..., 11

The orders are shown in detail in Table I (Appendix D), which indicates the question type for each of the orders for the 20 lotteries. This randomisation of the order of the lotteries was done in order to test speed of learning differences between the different presentation formats as it allowed us to distinguish between changes in difficulty and changes in timing. For instance if question 4 had a lower number of correct answers than say question 17, it could be because question 17 is an easier question or that question 17 just happened to come later in the sequence of questions. By also having an order where question 17 came before question 4, we can then estimate both a learning effect and a difficulty level.

The questions were classified as easy, hard or swing. There were eight easy questions and the remaining were equally divided between hard and swing. A simple rule of the thumb was used in the classification of the questions.

**Easy** – A hypothetical participant faced with the options where one statement is <= (or <) mode and the other > (or >=) mode, would always get the correct answer if one assumed mode=median.

More specifically, the "Correct" choice is:

If the statements are "above" and in A the threshold temperature is > (>=) the forecast mode and in B it is < (<=) the mode, the choice is B.

If the statements are "above" and in A the threshold temperature is <= (<) the forecast mode and in B it is > (>=) the mode, the choice is A.

If the statements are "below" and in A the threshold temperature is > (>=) the forecast mode and in B it is <= (<) the mode, the choice is A.

If the statements are "below" and in A the threshold temperature is <= (<) the forecast mode and in B it is > (>=) the mode, the choice is B.

For example question 1 (Figures 2 (a), (b) and (c)):

Statement A – The maximum temperature on Monday is **above** 6 deg. C

Statement B – The maximum temperature on Tuesday is **above** 0 deg. C

The mode for Monday is 5 deg. C and the mode for Tuesday is 1 deg. C, hence statement A > mode and statement B < mode therefore, if one assumes mode=median, the "correct"/most probable choice is B.

**Swing**– A hypothetical participant with just the point forecast assuming mode=median and same uncertainty at all forecasts (takes the distance from the point forecast as same deviation) would result in that participant choosing a different option compared to those with uncertainty information. Question 4 in Figure 1 was a swing question.

**Hard**– These were questions that were neither easy nor swing.

At the end of the experiments, participants were asked to fill out a brief questionnaire (Appendix B) and were then paid their total earnings in addition to a show up fee of £3. Most of the participants (44%) were from the business school whilst 37% were from humanities, the rest were science/engineering majors. Slightly more than half (52%) were female. Table II (Appendix D) shows the number of participants by school, format and order.

***Data***

The temperatures that were used in the experiments were not actual forecasts; rather the temperatures were generated using synthetic means. The 'observations' (answers/"actual" temperatures for each day) were produced using the triangular distribution with the peak at the stated "most likely" value and the tails beyond the stated "High range" and "Low range", to account for the 1/10 probability of observations falling outside the forecast range. The triangular distribution is a continuous probability distribution used when there is limited sample data or when the underlying probability is unknown (Kotz, 2004). It has three parameters; minimum, maximum and the mode/most likely value. The distribution is used in project management (e.g., Back et al., 1999; Larham, 2010), risk analysis (Johnson, 1997) and business decision making. The distribution is defined on the range $x \in [a, c]$ with the probability density function:[2]

$$f(x) = \begin{cases} \dfrac{2(x-a)}{(b-a)(c-a)} & \text{for} \quad a \le x \le b \\ \dfrac{2(c-x)}{(c-a)(c-b)} & \text{for} \quad b < x \le c \end{cases} \tag{1}$$

[2] Evans, M., Hastings, N. and Peacock, B., 2000. Statistical Distributions. Wiley-Interscience, New York , pg 187-188

And cumulative distribution functions:

$$F(x)=\begin{cases}\dfrac{(x-a)^2}{(b-a)(c-a)} & for \quad a\le x\le b\\[2ex] 1-\dfrac{(b-x)^2}{(b-a)(b-c)} & for \quad b< x\le c\end{cases} \tag{2}$$

Where $b\in[a,c]$ = mode or most likely, $a$ = minimum value, $c$ = maximum value

A diagram illustrating the PDF of the triangular distribtion[3] is shown in Figure 3.

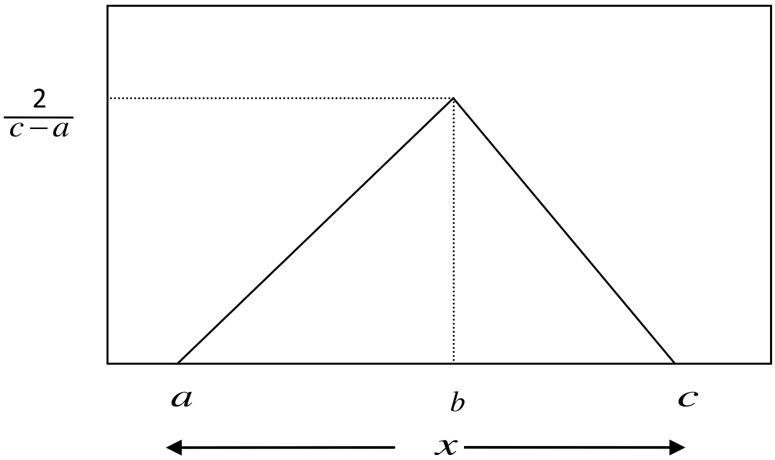



Figure 3: Illustration of the triangular distribution probability density function, see Hesse (2000)

## 3. Results

*Result 1: Participants who were provided with uncertainty information made better decisions than those without any uncertainty information.*

The average earnings for the three groups are summarised in Table III. Participants who were provided with uncertainty information (Groups B and C) earned more than those who did not have the extra information (Group A). This was statistically significant at the 1% level (t=8.1439 between Groups A/B and t=9.0582 between Groups A/C) whilst there are no significant differences in average earnings between B and C (t=1.0434). On average,

[3] The triangular distribution does not assume mode=median. This assumption may not always be true for real forecast distributions, but in some instances, it affects how participants might have interpreted the uncertainty information

participants who were provided with format B or C outperformed those with Format A regardless of the gender; school or order in which the questions were presented except for the non-swing questions (Table IV). Since 70% of the questions were non-swing most of Format A participants got them "correct". The "correct" answer was the more probable of the two statements that the participants were given as options.

Table III: Number of participants by treatment group and average earnings

| | Number of participants | Average earnings[1] | Std dev |
|---|---|---|---|
| Group A | 87 | £6.33 | 0.61 |
| Group B | 97 | £7.30 | 0.95 |
| Group C | 105 | £7.45 | 1.00 |

[1]these were average earnings not including the £3 show up fee

Table IV: Summary statistics by school, gender, order, whether or not it was a swing question and the associated Pearson chi squared values and probability. Chi square results indicate significant relationship between format and all categories shown except for order 4 participants.

| | Format A (% correct) | Format B (% correct) | Format C (% correct) | $Chi^2$ | p |
|---|---|---|---|---|---|
| Business | 64.9 | 80.0 | 75.6 | 41.59 | 0.000 |
| Humanities | 66.4 | 76.1 | 74.3 | 14.18 | 0.001 |
| Sciences | 67.6 | 77.5 | - | 13.80 | 0.000 |
| Female | 64.9 | 75.0 | 72.5 | 36.88 | 0.000 |
| Male | 67.6 | 80.2 | 78.2 | 31.79 | 0.000 |
| Swing questions | 18.0 | 57.9 | 59.7 | 244.74 | 0.000 |
| Non swing questions | 86.9 | 86.0 | 81.7 | 16.40 | 0.000 |
| Order 1 | 66.6 | 79.2 | 75.0 | 22.25 | 0.000 |
| Order 2 | 67.4 | 79.6 | 78.0 | 22.30 | 0.000 |
| Order 3 | 62.0 | 76.7 | 75.5 | 28.04 | 0.000 |
| Order 4 | 70.0 | 75.3 | 70.7 | 3.90 | 0.142 |
| Overall | 66.2 | 77.6 | 75.1 | 66.24 | 0.000 |

Group C participants, who were shown a graph with uncertainty information, did slightly better than those presented with the table with uncertainty information (Group B) for the swing questions on average. The difference between Group A and either Group B or C participants is greatest for the swing questions: (39.9% for B and 41.7% for C). Overall, on average participants chose the most probable outcome approximately 66.2%, 77.6% and

75.1% of the time for Group A, B and C respectively. Chi squared tests indicate significant association between the format and the different categories except when participants are shown questions in order 4. This is possibly from a random sampling effect. There is no significant association between Format B/A for the non-swing questions (chi=0.3997 and p=0.527). Table V summarises the percentage of participants who answered each question correctly, type of question, probability of statement A or B being correct and which statement was actually true. The number of questions with the highest percentage correct is almost equally divided between the 3 groups. For all the swing questions, those with uncertainty information did better than those without. Chi squared tests show that there is a significant relationship at the 1% level between the format shown and question type for all the swing questions and also questions 1, 5 and 12 which were easy questions. At the 5% level there is significant association between format and questions 8 and 9. Further statistical analysis is done to determine the degree and direction of this association.

We analysed if there was an increase in the percentage of participants choosing the most likely outcome as the experiment progressed for the different question types by order (refer to Table I for classification for each order). Results are summarised in Table VI. For example if we look at Order 1, 31% of the of the participants chose the most likely outcome on their first swing question and on the 5th swing question, 73.8% participants chose the correct statement. Similarly, 84% of the participants shown Order 4 get the 1st easy question correct and 81%, 92%, 94% ,66% and 95% choose the correct outcome in the 2nd , 3rd,..., and so on easy question. From the analysis it is not easy or sufficient to conclude whether or not there was a learning effect, hence the need to do further statistical analysis using regression in order to find out whether or not there was a significant increase in the probability of participants choosing the correct answer or whether they took less time as the experiment progresses by format and order. Choosing the most probable outcome most likely also depended on the question which was asked before.

Table V: Summary showing the percentage of participants who answered each question correctly, Pearson chi squared and its probability, the probabilities of statements A or B being correct, which of the statements was true, classification of the questions and whether or not the earlier day was the most probable outcome. Chi square results show significant association between format and all swing questions at the 1% level.

| Question | Format A | Format B | Format C | $Chi^2$ | p | Probability | | Actual | Class | Early correct |
|---|---|---|---|---|---|---|---|---|---|---|
| | % Correct | % Correct | % Correct | | | A | B | | | |
| 1 | 97.7 | 72.2 | 62.9 | 33.61 | 0.000 | 0.56 | 0.78 | 1 | easy | N |
| 2 | 92.0 | 92.8 | 89.5 | 0.74 | 0.692 | 0.05 | 0.55 | 2 | easy | N |
| 3 | 9.2 | 39.2 | 23.8 | 22.30 | 0.000 | 0.78 | 0.69 | both | swing | Y |
| 4 | 19.5 | 63.9 | 73.3 | 61.23 | 0.000 | 0.26 | 0.38 | 2 | swing | N |
| 5 | 62.1 | 92.8 | 87.6 | 32.80 | 0.000 | 0.75 | 0.40 | 1 | easy | Y |
| 6 | 4.6 | 60.8 | 59.0 | 75.84 | 0.000 | 0.30 | 0.38 | 2 | swing | N |
| 7 | 92.0 | 85.6 | 89.5 | 1.96 | 0.375 | 0.78 | 0.92 | both | hard | N |
| 8 | 88.5 | 79.4 | 74.3 | 6.15 | 0.046 | 0.39 | 0.22 | 1 | easy | Y |
| 9 | 82.8 | 85.6 | 72.4 | 6.09 | 0.048 | 0.87 | 0.61 | 1 | hard | Y |
| 10 | 88.5 | 87.6 | 87.6 | 0.04 | 0.978 | 0.19 | 0.64 | 2 | easy | N |
| 11 | 34.5 | 53.6 | 71.4 | 26.20 | 0.000 | 0.19 | 0.30 | 2 | swing | N |
| 12 | 80.5 | 66.0 | 49.5 | 20.02 | 0.000 | 0.58 | 0.49 | 2 | easy | Y |
| 13 | 85.1 | 84.5 | 76.2 | 3.30 | 0.192 | 0.58 | 0.38 | 2 | easy | Y |
| 14 | 89.7 | 90.7 | 91.4 | 0.18 | 0.915 | 0.70 | 0.89 | both | hard | N |
| 15 | 89.7 | 88.7 | 90.5 | 0.18 | 0.914 | 0.95 | 0.57 | 1 | hard | Y |
| 16 | 25.3 | 75.3 | 85.7 | 83.43 | 0.000 | 0.05 | 0.19 | neither | swing | N |
| 17 | 80.5 | 85.6 | 85.7 | 1.22 | 0.543 | 0.60 | 0.78 | 2 | hard | N |
| 18 | 95.4 | 95.9 | 91.4 | 2.17 | 0.337 | 0.47 | 0.84 | 2 | hard | N |
| 19 | 92.0 | 96.9 | 95.2 | 2.35 | 0.309 | 0.50 | 0.78 | 2 | easy | N |
| 20 | 14.9 | 54.6 | 44.8 | 32.57 | 0.000 | 0.86 | 0.72 | both | swing | Y |

Table VI: Summary showing the percentage of participants who answered each round correctly by order. Chi square results show significant association between the order in which the questions were asked and round for all questions at the 1% level.

| Round | Order 1 (n=84) | Order 2 (n=76) | Order 3 (n=65) | Order 4 (n=64) | $Chi^2$ | p |
|---|---|---|---|---|---|---|
| | % Correct | % Correct | % Correct | % Correct | | |
| 1 | 89.3 | 46.1 | 41.5 | 84.4 | 60.81 | 0.000 |
| 2 | 89.3 | 94.7 | 58.5 | 82.8 | 36.34 | 0.000 |
| 3 | 31.0 | 92.1 | 78.5 | 81.3 | 82.82 | 0.000 |
| 4 | 48.8 | 84.2 | 86.2 | 84.4 | 40.77 | 0.000 |
| 5 | 75.0 | 60.5 | 87.7 | 34.4 | 45.49 | 0.000 |
| 6 | 42.9 | 89.5 | 60.0 | 92.2 | 61.17 | 0.000 |
| 7 | 85.7 | 93.4 | 84.6 | 51.6 | 43.37 | 0.000 |
| 8 | 75.0 | 86.8 | 95.4 | 23.4 | 99.57 | 0.000 |
| 9 | 79.8 | 68.4 | 95.4 | 93.8 | 24.80 | 0.000 |
| 10 | 84.5 | 61.8 | 29.2 | 65.6 | 48.04 | 0.000 |
| 11 | 59.5 | 92.1 | 75.4 | 42.2 | 44.16 | 0.000 |
| 12 | 60.7 | 77.6 | 89.2 | 95.3 | 31.64 | 0.000 |
| 13 | 82.1 | 88.2 | 18.5 | 93.8 | 121.49 | 0.000 |
| 14 | 91.7 | 94.7 | 52.3 | 87.5 | 55.41 | 0.000 |
| 15 | 90.5 | 42.1 | 81.5 | 59.4 | 51.23 | 0.000 |
| 16 | 73.8 | 80.3 | 53.9 | 90.6 | 24.85 | 0.000 |
| 17 | 81.0 | 63.2 | 90.8 | 90.6 | 23.23 | 0.000 |
| 18 | 95.2 | 23.7 | 76.9 | 78.1 | 103.48 | 0.000 |
| 19 | 94.1 | 93.4 | 80.0 | 70.3 | 22.25 | 0.000 |
| 20 | 38.1 | 72.4 | 90.8 | 51.6 | 49.60 | 0.000 |

Statistical analysis using a probit regression model estimated the determinants of choosing the most probable outcome. In our study, we were interested in the factors (predictors) that determine whether or not a participant chose the 'correct' (most likely) outcome. In order to do this a probit regression model was used. There were two outcomes (dependent variable) that were possible (participant chooses the 'correct' / most likely outcome OR participant chooses a 'wrong' outcome). Hence the dependent variable (y) is binary i.e. y=1 if participant chooses the most likely outcome and y=0 otherwise. Some of the predictor variables ($x$) were, whether or not a participant was provided with uncertainty information (i.e. Format A, B or C), gender of the participant, whether or not the question was swing, how often they check the weather and so forth. The objective of using a probit model was to find the best fitting model to describe the relationship between the dependent variable and the predictor variables. This model was chosen instead of the conventional OLS because of our binary dependent variable. Hence, the probability that a participant will choose the most likely outcome given a set of predictor variables is given by: $\Pr(y = 1|x) = F = \phi(x\beta)$, where $\phi$ is cumulative distribution function of the standard normal distribution, $x$ is a set of explanatory variables and $\beta$ are parameters to be estimated (coefficients in table VII). The results of fitting a probit model to the data are shown in Table VII. All the predictor variables were dummy variables except for 'time', 'round number' and all the round number interaction terms. Table VIII (Appendix D) describes the variables used in the econometric analyses.

Marginal effects were also computed and are shown as the percentage change in probability in Table VII. The probit results only showed the relationship between whether or not a participant would choose the most likely outcome and the predictor variables but not the magnitude of the negative/positive effects. Marginal effects are evaluated by taking the derivative of the probability of choosing the probable outcome associated with a certain predictor ($dF/dx$). For example, on average, the probability of a male participant choosing the most probable outcome was 2.2% higher than that of the female participants at the 10% level of significance. Two control variables were included as dummy variables in the probit regression model. The variable, Test question Dummy' equals one for those participants that were shown test questions worded as the 'most likely' maximum temperature for Treatments B and C instead of the maximum temperature that is 'expected'. The test questions were asked at the beginning of the experiment. The other control variable 'Format B with description' is equal to one for Group B participants with the test questions worded as '*most likely'* and were given tables mentioning that the maximum temperature ranges fall within the indicated range 9 out of 10 times, and zero for Group B participants who were shown tables that did not mention that the maximum temperature ranges fall within the indicated range 9 out of 10 times (i.e. Figure 1(b) or 2(b) without the 'product description') and all the other participants.

*Result 2: Further analysis using probit regression model indicates that participants who were provided with uncertainty information were more likely to choose the most probable outcome compared to those who were shown the table without uncertainty information in general and in the swing questions.*

Results from the probit results indicate that 'Swing question, 'swing question & Format A', Format B and Format C were significant determinants of choosing the most likely outcome. If a question was a swing question, the probability of participants picking up the most probable outcome was reduced by 32.3% on average and if participants did not have uncertainty information (Format A), they were on average 73.7% less likely to get the swing questions correct at the 1% level of significance. For those students provided with the graph with uncertainty information, the probability of them choosing the correct outcome was increased by 10.7% on average whilst participants who were provided with the table with uncertainty information were 7.4% more likely to opt for the most probable outcome on average at the 1% and 5% level of significance respectively. The graph was a stronger predictor compared to the table.

*Result 3: For some of the questions, providing uncertainty information was not always useful.*

For some of the questions, participants who were provided with uncertainty information were less likely to choose the correct outcome if one of the options had a bigger area (length) between the high/low range and the asked temperature than the other one. This would not necessarily mean however that this was the most probable outcome. Providing uncertainty information for questions 1, 12 and 13 reduced the likelihood of choosing the correct outcome by 14.3% and 19.2% for Group B and Group C participants respectively (for example in question 1, which is an easy question, shown in Figure 2, participants with uncertainty information (28% and 37% for format B and C respectively) choose statement A as the most likely outcome because of the large area above 6 deg. C on Monday but the most probable outcome would be statement B.

*Result 4: Participants were more inclined to choose the option with the later date*

This result is consistent with the Roulston and Kaplan 2009 study. Participants were more inclined towards choosing the option with the later date (i.e. statement B). 'Early correct' means that the predictor was equal to one if the option with the earlier date (e.g. Wednesday) was the most likely as opposed to the later one (e.g. Friday).

Table VII: Results from probit regression model predicting the probability of selecting the most likely outcome as a function of predictor variables. The p-value indicates the statistical significance of each of the predictor variables.[4]

| Predictor (*see Table VIII in Appendix D for full descriptions*) | Coefficient | P | % change in probability |
|---|---|---|---|
| Round number | 0.005 | 0.195 | 0.14 |
| Swing question | -0.961*** | 0.000 | -32.28 |
| Swing question & Format A | -1.129*** | 0.000 | -41.40 |
| Swing question & Format C | 0.151 | 0.144 | 4.40 |
| Male | 0.072* | 0.084 | 2.20 |
| English is first language | 0.052 | 0.316 | 1.61 |
| Checks internet for weather forecast | 0.004 | 0.963 | 0.13 |
| Length | -0.416*** | 0.000 | -14.25 |
| Area | -0.546*** | 0.000 | -19.16 |
| Checks weather at least every 2-3 days | -0.053 | 0.211 | -1.62 |
| Sample question mistake | -0.361*** | 0.000 | -12.21 |
| Die question mistake | -0.197*** | 0.000 | -6.20 |
| Format B | 0.250* | 0.067 | 7.39 |
| Format C | 0.363*** | 0.007 | 10.67 |
| Order 1 | 0.223** | 0.043 | 6.57 |
| Order 1 & Format B | -0.056 | 0.689 | -1.74 |
| Order 1 & Format C | -0.284** | 0.036 | -9.33 |
| Order 2 | 0.182 | 0.116 | 5.39 |
| Order 2 & Format B | -0.058 | 0.693 | -1.79 |
| Order 2 & Format C | -0.177 | 0.207 | -5.68 |
| Order 4 | 0.302** | 0.026 | 8.63 |
| Order 4 & Format B | -0.397** | 0.013 | -13.38 |
| Order 4 & Format C | -0.581*** | 0.000 | -20.40 |
| Early day correct | -0.397*** | 0.000 | -12.42 |
| Early day correct & Format A | -0.126 | 0.238 | -3.99 |
| Early day correct & Format C | -0.297*** | 0.002 | -9.73 |
| Early day correct & Order 1 | 0.040 | 0.718 | 1.22 |
| Early day correct & Order 2 | 0.145 | 0.207 | 4.24 |
| Early day correct & Order 4 | 0.314*** | 0.008 | 8.64 |
| Above & certain | -0.160** | 0.039 | -5.06 |
| Above & uncertainty | -0.316*** | 0.000 | -9.87 |
| Test question dummy | 0.032 | 0.643 | 0.97 |
| Business | 0.008 | 0.909 | 0.24 |
| Humanities | -0.012 | 0.852 | -0.37 |
| Format B with description | -0.045 | 0.661 | -1.39 |
| Hard question | 0.086 | 0.149 | 2.60 |
| Constant | 1.166*** | 0.000 | 0.00 |

***, ** and * indicates statistical significance at the 1%, 5% and 10% level respectively.

[4] Order 3 was dropped as the default dummy group in the regression models.

Participants were on average, 12.4% less likely to choose the correct outcome whilst if they were presented with Format C, participants were on average 22.2% less likely to choose the most probable outcome if it was on an earlier date at the 1% level.

*Result 5: Male participants were more likely to choose the most probable outcome.*

On average, male participants were 2.2% more likely to choose the most likely outcome compared to their female counterparts at the 10% level of significance.

*Other probit results*

Participants who answered the sample question and probability question incorrectly were less likely to choose the most probable outcome. The participants were asked a basic probability question on the probability of a six appearing if a fair die was rolled twice, and participants who answered the die question incorrectly were on average 6.2% less likely to choose the correct statement. Participants who got at least two of the test questions wrong had their chance of choosing the correct option reduced by 12.2% on average. This may indicate that it is necessary for people to be able to read diagrams with weather forecast information in order to interpret them and make correct decisions. The way the participants interpreted the data and used it to make decisions based on the test questions is however beyond the scope of this study.

When the question asked for the maximum temperature above X deg. C instead of below, it reduced the probability of the participants choosing the most probable outcome, whether or not they had the uncertainty information; this was reduced more for participants who were provided with uncertainty information (Groups B & C) who were on average 9.9% less likely to choose the correct option compared to 5.1% for those without. Other predictor variables that were statistically significant at the 5% level were related to the order the questions were asked. . If participants were shown questions using order 4, the probability that they would choose the most likely outcome was increased by 8.6% on average than if the questions were presented in order 3, however if participants were shown a table with uncertainty information (Format B), they were 4.8% less likely to choose the correct outcome on average. Order 1 participants were 6.6% more likely to select the correct outcome; if their presentation format was a graph with uncertainty information (Format C), they were 2.8% less likely to choose the correct outcome, on average. There is need to however emphasize that the orders were randomly assigned and the low numbers of participants for some of the orders, particularly order 4 might account for the significant effects.

Whether or not participants checked the weather frequently (at least every 2-3 days) and whether or not they got weather forecasts from the internet did not have a statistically significant impact on performance. The control variables, 'test question dummy' and 'format B with description' were also not statistically significant. A second probit model run with the variables,' Round number squared', 'Round number & Format B', 'Round number & Format C', 'Round number squared & Format B' and 'Round number squared & Format C' to test possible learning effects for the two formats has same results for the model shown and none of them are significant predictors of choosing the most probable outcome (except 'Format B' becomes insignificant with p=0.954 and 'Format C' has less predictive power, p=0.074).

*Result 6: Time analysis*

*There was a learning effect as the experiment progressed. Participants who were shown the graph with uncertainty information took on average less response time compared to those who were shown a table with uncertainty information.*

The total average response and review time for all the 20 'lotteries ranged from 26.9sec to 63.1sec. On each question, the average response time to answer each question was 30.8sec and participants took on average 7.6sec to review the results of each lottery. Figure 5 shows a general decrease in the time participants took to respond to the questions as the experiment progresses. Figure 6(a), (b), (c) and (d) shows the average response times by group and question order. There is a general decrease in response time for the 4 orders as the experiment progresses. The differences in order 4 could be due to the small sample size for Group A participants. Figure 7 (Appendix D) displays the average response times participants took to answer the easy, swing and hard questions for the four different orders. Group B participants who were provided with a table with uncertainty information took more time despite the question type compared to Group A (no uncertainty information) and Group C participants (graph with uncertainty information) except when the question order is 4, where Group C participants take more time on average. In most cases, participants on average took more time on the easy questions.

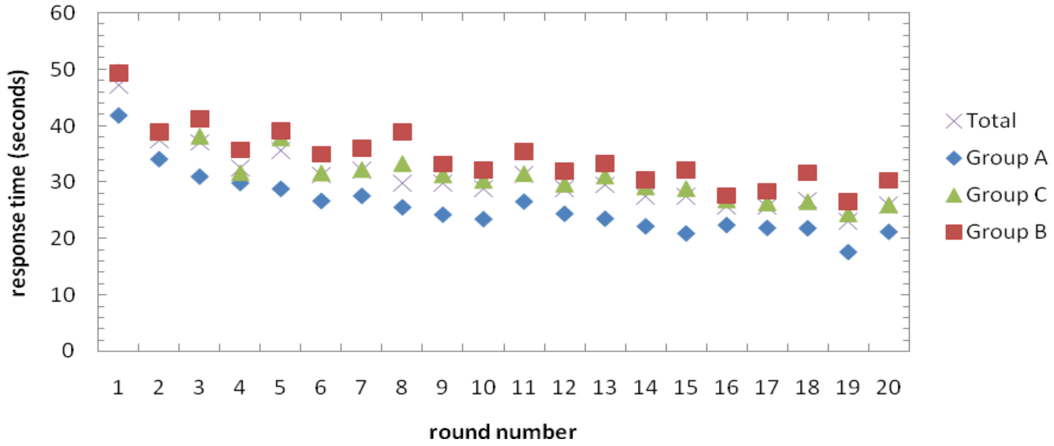



Figure 5: Average response time participants took for each round. There is a general decrease in time spent per question as the experiment progresses. Group A participants were faster than the other two groups whilst there isn't much difference between Groups B and C, although Group C is slightly faster. Statistical analysis will be done to test if the difference between Groups B and C is significant.

When the median response time (Figure 8) instead of the average is used, results show that participants took more time on the swing questions compared to the other question types despite the format shown or order except for Group B participants with order 1, where the easy took more time. For the easy and hard questions, participants took almost the same median time for Order 2, 3 and 4.

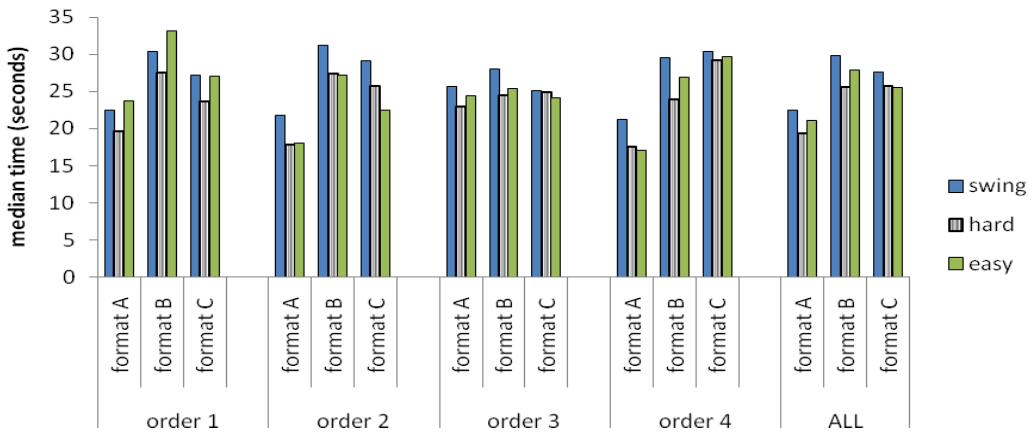



Figure 8: Median response time for each of the groups by order and question type

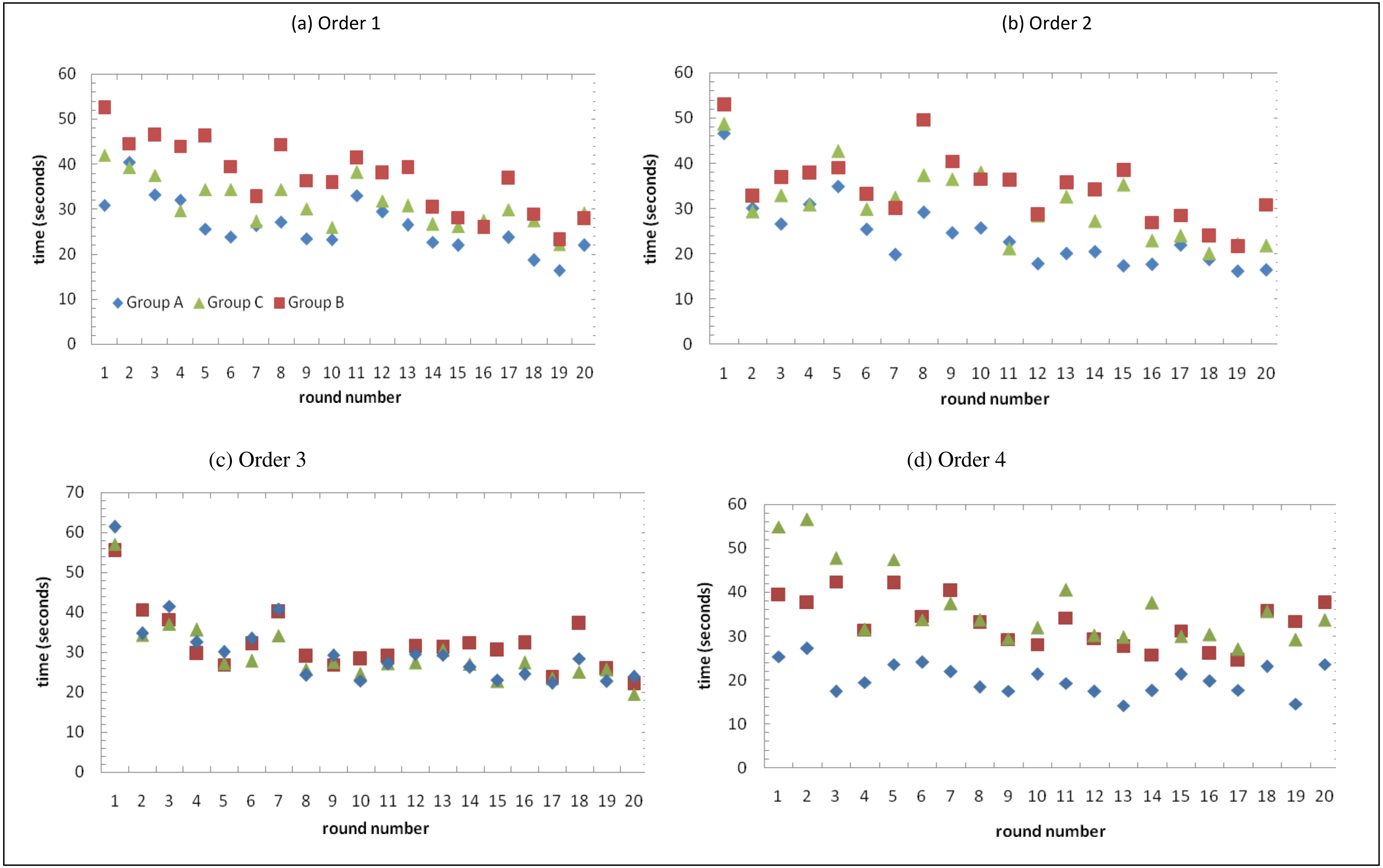



Figure 6(a), (b), (c) and (d): Average response times taken by participants to answer questions in each round for the 3 groups differentiated by question order.

Statistical analysis was done using multiple linear regressions (MLR) to assess the determinants of time participants who were provided with uncertainty information took to respond to questions. MLR models the relationship between a dependent variable which in this case is response time and a set of independent variables. Results of the MLR are shown in Table XI. The time regression model was basically done to measure the differences between Format B and Format C with the null hypothesis that participants who were shown the table with uncertainty information (Format B) took more response time compared to the participants who were shown the graph with uncertainty information (Format C), hence only data from participants who were shown the two formats was used. The regression model can be written as:

$$responsetime = \beta_0 + \beta_1\chi_1 + \beta_2\chi_2 + ... + \beta_k\chi_k + \mu_i$$

where, $\beta_0$ is the intercept, $\beta_i$ to $\beta_k$ are the coefficients on the *k* independent variables, $\chi_i$ are the independent variables that affect *responsetime* (for example whether or not English is first language) and $\mu_i$ is the error term which contains other explanatory variables which are not included in the model. The description of the variables is the same as shown in Table VIII (BUT note that as mentioned previously the data used in the MLR was from participants who were shown format B and C only hence for the dummy variables description; 0 = for Format C participants for example 'Format B' means that 1= participant was presented with a table with uncertainty information, 0= participant was presented with a graph with uncertainty information whilst 'Swing and Format B' means that 1= question was swing & participants were shown a table with uncertainty information, 0= question was swing participants & were shown Format C graph. Format C participants were the base group.

Thirteen of the explanatory variables are significant determinants of response time at the 1% level. As the experiment progressed, participants were taking 1.7 seconds less from round to round at the 1% level of significance, all other things constant. This might indicate a learning effect. The effect is negative until the last round with a decrease in time of 0.05s. There is a positive relationship between whether or not a question is swing and the response time participants took. If a question was swing, participants would take 4.3 seconds more compared to the other questions, all things being constant. Participants whose first language was English took 3.5s less time holding all the other factors constant. This result is expected, since the experiment was conducted in English; hence, native English speakers would take less time compared to participants whose first language is not English. Participants in Group B who made their decisions based on the length above the temperature asked on the options given (i.e. chose the option with the largest difference between the high/low range and asked temperature) took 5.9s more to respond to questions, all things constant. Order 4 is the only one of the question orders that is significant. Participants shown questions in order 4 took 6s more to respond to question and for participants who were shown the table with uncertainty information, the response time is reduced to 0.6s, all other things constant.

Table XI: Results of linear regression analysis with response time as the dependent variable

| Independent variable | Coefficient | T-value | P-value |
|---|---|---|---|
| Round number | -1.67 | -7.21*** | 0.000 |
| Swing question | 4.27 | 3.97*** | 0.000 |
| Swing & Format B | 0.42 | 0.29 | 0.772 |
| Male | -0.29 | -0.42 | 0.676 |
| English is first language | -3.51 | -4.20*** | 0.000 |
| Checks internet for weather forecast | -4.27 | -3.18*** | 0.001 |
| Length | 5.89 | 3.98*** | 0.000 |
| Area | 2.95 | 2.06** | 0.039 |
| Checks weather at least every 2-3 days | 3.14 | 4.69*** | 0.000 |
| Sample question mistake | -2.76 | -2.00** | 0.045 |
| Die question mistake | 1.98 | 2.62*** | 0.009 |
| Order 1 | 1.64 | 1.31 | 0.191 |
| Order 1 & Format B | 2.88 | 1.56 | 0.119 |
| Order 2 | 1.24 | 0.98 | 0.329 |
| Order 2 & Format B | 0.92 | 0.49 | 0.621 |
| Order 4 | 6.05 | 4.37*** | 0.000 |
| Order 4 & Format B | -5.49 | -2.83*** | 0.005 |
| Early day correct | 1.09 | 1.18 | 0.238 |
| Early day correct & Format B | 0.18 | 0.13 | 0.894 |
| Test question dummy | 5.37 | 5.59*** | 0.000 |
| Science | 1.72 | 1.21 | 0.226 |
| Humanities | -2.99 | -3.44*** | 0.001 |
| Format B with description | -5.28 | -3.56*** | 0.000 |
| Easy question | -0.28 | -0.30 | 0.764 |
| Round number squared | 0.04 | 3.78*** | 0.000 |
| Format B | 3.61 | 1.89* | 0.059 |
| Above | -2.07 | -2.21** | 0.027 |
| Above & Format B | 0.25 | 0.19 | 0.846 |
| Constant | 44.46 | 19.79*** | 0.000 |

***, ** and * indicates statistical significance at the 1%, 5% and 10% level respectively.

The weather variables, 'Checks internet for weather forecast' and 'Checks weather at least every 2-3 days' were significant at the 1% level. The former show a decrease in the time participants take to respond of 4.3s whilst participants who indicated the latter took 3.1s more to respond, all other things constant. Students studying humanities took 3.0s less to respond to questions. Those participants who were shown the new test questions and those who got the probability question wrong took 5.4s and 2s more respectively. Students who were shown the table with uncertainty information (including the product description stating that the temperatures fell within the range 9 out of 10 times) took 5.3s less than those who were shown the graph with uncertainty information. However all the participants who were shown the table with uncertainty information took 3.6s more compared to those shown the graph with uncertainty information at the 10% level of significance, everything else constant.

The explanatory variables that are significant at the 5% level of significance are 'Area', 'Test question mistake', and 'Above'. Holding all other things constant, participants who got at least two of the test questions wrong took 2.8s less to respond whilst if the questions asked for the temperature 'above' X$^{o}$C rather than 'below', participants took 2.1s less time to respond. Participants in Group C who made their decisions based on the area above the temperature asked on the options given (i.e. chose the option with the largest difference between the high/low range and asked temperature) took 2.9s more to respond to questions, all things constant.

## 4. Conclusion

Experimental economics methodology was used to assess the impact of providing probabilistic weather information on decision making using different presentation forecasts to undergraduate students at the University of Exeter.

As in the previous study by (Roulston and Kaplan, 2009), participants who were provided with uncertainty information made better decisions compared to those without any uncertainty information. Their study used fan charts based on a 5-day forecast whilst this study compared and contrasted bar graphs and tables. Initial descriptive analysis shows an improvement in decision making for participants who were shown uncertainty information (Format B/C) despite the gender, academic department, or order in which questions were asked. Results indicate that participants who were provided with uncertainty information were able to use this information to make better decisions, especially for the swing questions as those without uncertainty information could not choose the correct answer on these questions. (Swing questions are those where a hypothetical participant with just the point forecast, assuming mode=median, would result in that participant choosing a different option compared to those provided with uncertainty information.) There is not much difference for some of the non-swing questions between the three formats, with those with uncertainty information doing worse in some of the questions. However, overall, participants provided with uncertainty information did better.

Participants who were provided with a graph with uncertainty information did slightly better than those participants who were provided with a table with uncertainty information on the swing questions. Further analysis using statistical means indicates that both the table and the graph with uncertainty information were significant determinants of choosing the most probable outcome, though the graph is a stronger predictor compared to the table.

Providing uncertainty information was not always useful for some of the questions. In some instances, the participants with the graph/table with uncertainty information were making their decisions by choosing the day with the largest area between the asked temperature and high/low range, which might not have been the most probable outcome, hence it is possible that some did not understand or could not use the uncertainty information correctly. This was indicated by the highly significant ‘length’ and ‘area’ variables in the probit models.

There is a general decrease in time spent per question as the experiment progresses for all the 3 groups. Statistical analysis shows a significant decrease in the time participants took to respond to questions as the experiment progressed, indicating a learning effect. This is useful as it indicates that interpretation of a particular presentation of forecasts becomes easier with familiarity. Participants who were shown the graph with uncertainty information took on average less response time compared to those who were shown a table with uncertainty information. The difference however was small between the two formats and it cannot be used to conclude on interpretation or how subjects were using the information.

Assessing how weather information is disseminated to users is useful. Including information on uncertainty better represents the capability of the forecasts and provides the potential for better decision making, but information has to be communicated effectively in a format that users can interpret, use efficiently and avoid poor decisions. Results from our analysis can be used by the Met Office to help them decide which format to use to best disseminate weather information to the public and other partners, and to determine the value of presenting probabilistic information. Results also add to the knowledge on weather forecast interpretation and decision making when provided with uncertainty information presented in a bar graph or table format. Research using the general public is also essential (a study is currently underway at the Met Office using a weather game with the UK general public). A follow up study using these formats can be used to do a study on perhaps real decision making, on simple decisions such as whether or not to carry an umbrella or go out to a picnic, farmers deciding when to plant, whether or not to go ahead with a sport match; to national decisions such as deciding whether to evacuate people or fly a plane. Results can potentially help various sectors that use weather information which include agriculture, aviation, sports, energy, as well as policy makers and the general public. Other potential applications of the study include pensions giving risk advice, brokers giving investment advice, and government displaying economic forecasts.

*Acknowledgements:* We would like to express our gratitude to Tim Miller, Sarunas Girdenas, Stephen Pearson for research assistance and the participants for taking part in the experiments. The experiments were conducted at the Finance and Economics Experimental Laboratory (FEELE) at the University of Exeter. We also wish to thank the Met Office for funding this research.

## Appendix A: Instructions

'You are about to participate in an experiment involving the interpretation of weather forecasts. If you follow the instructions carefully and make wise decisions, you may earn a significant amount of money. Your earnings will depend on your decisions. Participants in this experiment do not interact with one another, so your earnings do NOT depend on the decisions of the other participants. All of your decisions will remain anonymous and will be collected through a computer network. Your decisions are to be made at the computer at which you are seated. Your total earnings from the experiment will be paid to you, in cash, at the end of the experiment.

Please turn off your mobile phone and do NOT attempt to communicate with the other participants. If you have any questions, please RAISE YOUR HAND and someone will come and help you. It is important that you understand the instructions. Misunderstandings may result in lower earnings.

The experiment consists of 20 repeated rounds. In each round you will be shown a graphic of the predicted maximum temperature over the course of the next few days, similar to the one below. You will also see two statement about the future weather, called Statement A and Statement B, for example:

Statement A – The maximum temperature on Saturday is above 12 °C
Statement B – The maximum temperature on Tuesday is above 15 °C

Statements A and B may or may not be true. In other words, neither statement may be true, both statements may be true or only one of the two may be true. The statements relate to the ACTUAL temperature. Your task is to study the graph, which shows the FORECAST temperature, and work out which statement is MORE LIKELY to be true.

You will begin each round by looking at the graphic and the statements and then choosing ONE of the two statements, either Statement A or Statement B. After you have chosen, you will be told the actual maximum temperatures on the days in question and therefore whether or not each statement is true. If you chosen statement is true, you will receive a payoff of 1 Feele token, otherwise you will not receive a payoff. Feele tokens will be converted into cash at the end of the experiment, at a rate of 50 pence per token. You will also receive a show-up fee of £3.00 for participating in this experiment.

There are NO trial rounds, so when you start 'Round 1', you will be playing for real money. Before you do so, however, you will be asked to answer some multiple choice test questions. The answers you give to these questions do NOT affect your payment; the idea is for you to get some practice reading graphs.

If you have any questions, either now or later on, please raise your hand and someone will come and help you.'

## Appendix B: Supplementary questions

1. What is your gender?
   (F) female; (M) male
2. Is English your first language

(Y) yes; (N) no
3. Which of the following sources of information do you consult when you want to find out the weather forecast?
   (a) Internet; (b) TV; (c) Radio; (d) Newspaper; (e) Ask someone else
4. How frequently do you look at the weather forecast?
   (a) Never or hardly ever; (b) Weekly; (c) Every 2 or 3 days; (d) Daily; (e) More than once a day
5. If a fair die is rolled twice, what is the probability that a six will appear on both occasions?
   (a) 0; (b) 1/6; (c) 1/12; (d) 1/36; (e) None of the above
6. What strategies did you use to make your decisions?

**Appendix C: Test question example**

Version 1: What is the maximum temperature that is *expected* on Saturday?
Version 2: What is the maximum temperature that is *most likely* on Saturday?

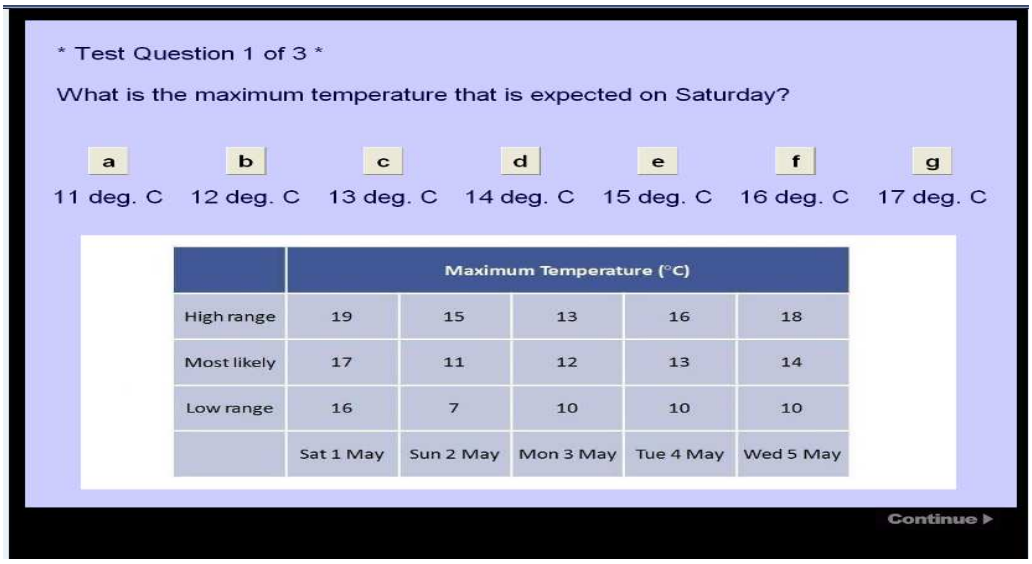



Table X: Number of participants for the two different versions of test questions used. Test questions were either worded as ‘expected’ or as ‘most likely’

| Order | Format A | Format B | | Format C | |
|---|---|---|---|---|---|
| | **‘expected’** | **‘expected’** | **‘most likely’** | **‘expected’** | **‘most likely’** |
| 1 | 28 | 14 | 10 | 22 | 10 |
| 2 | 23 | 13 | 10 | 20 | 10 |
| 3 | 22 | 12 | 9 | 15 | 7 |
| 4 | 14 | 20 | 9 | 13 | 8 |
| Total | **87** | **59** | **38** | **70** | **35** |

**Appendix D: Tables and Figures**

Table I: Summary of the 4 orders that were used and the type of question for the 20 rounds (lotteries)

| Round | Order 1 | | Order 2 | | Order 3 | | Order 4 | |
|---|---|---|---|---|---|---|---|---|
| | Class | Question number | Class | Question number | Class | Question number | Class | Question number |
| 1 | easy | 1 | swing | 20 | swing | 11 | easy | 10 |
| 2 | easy | 2 | easy | 19 | easy | 12 | hard | 9 |
| 3 | swing | 3 | hard | 18 | easy | 13 | easy | 8 |
| 4 | swing | 4 | hard | 17 | hard | 14 | hard | 7 |
| 5 | easy | 5 | swing | 16 | hard | 15 | swing | 6 |
| 6 | swing | 6 | hard | 15 | swing | 16 | easy | 5 |
| 7 | hard | 7 | hard | 14 | hard | 17 | swing | 4 |
| 8 | easy | 8 | easy | 13 | hard | 18 | swing | 3 |
| 9 | hard | 9 | easy | 12 | easy | 19 | easy | 2 |
| 10 | easy | 10 | swing | 11 | swing | 20 | easy | 1 |
| 11 | swing | 11 | easy | 10 | easy | 1 | swing | 20 |
| 12 | easy | 12 | hard | 9 | easy | 2 | easy | 19 |
| 13 | easy | 13 | easy | 8 | swing | 3 | hard | 18 |
| 14 | hard | 14 | hard | 7 | swing | 4 | hard | 17 |
| 15 | hard | 15 | swing | 6 | easy | 5 | swing | 16 |
| 16 | swing | 16 | easy | 5 | swing | 6 | hard | 15 |
| 17 | hard | 17 | swing | 4 | hard | 7 | hard | 14 |
| 18 | hard | 18 | swing | 3 | easy | 8 | easy | 13 |
| 19 | easy | 19 | easy | 2 | hard | 9 | easy | 12 |
| 20 | swing | 20 | easy | 1 | easy | 10 | swing | 11 |

Table II: Number of participants by school, format and order

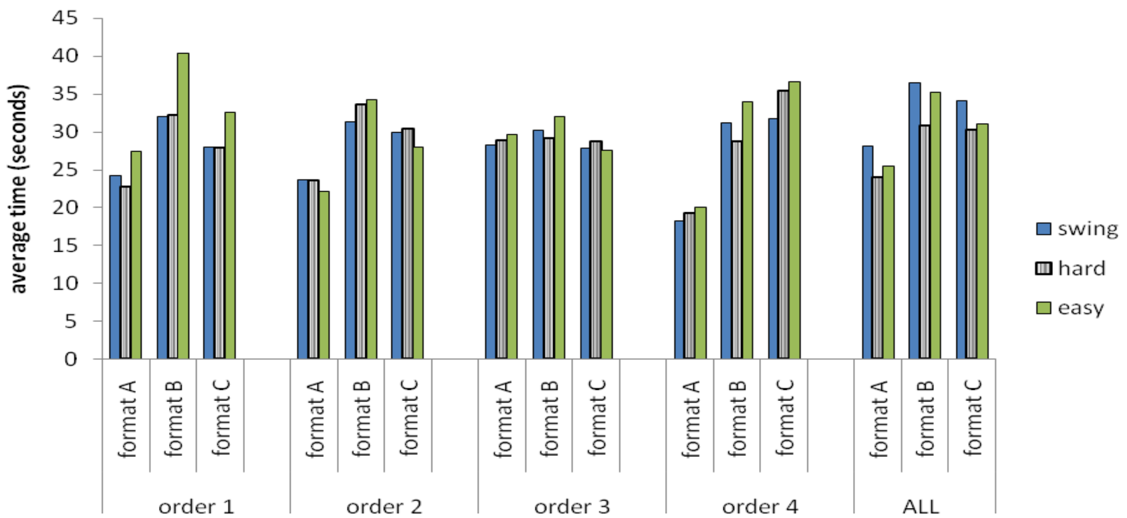
| | Order | Business | Humanities | Sciences | Total |
|---|---|---|---|---|---|
| Format A | 1 | 14 | 6 | 8 | 28 |
| | 2 | 9 | 6 | 8 | 23 |
| | 3 | 9 | 5 | 8 | 22 |
| | 4 | 3 | 4 | 7 | 14 |
| | | 35 | 21 | 31 | 87 |
| Format B | 1 | 5 | 12 | 7 | 24 |
| | 2 | 5 | 11 | 7 | 23 |
| | 3 | 5 | 9 | 7 | 21 |
| | 4 | 12 | 12 | 5 | 29 |
| | | 27 | 44 | 26 | 97 |
| Format C | 1 | 21 | 11 | - | 32 |
| | 2 | 19 | 11 | - | 30 |
| | 3 | 13 | 9 | - | 22 |
| | 4 | 11 | 10 | - | 21 |
| | | 64 | 41 | | 105 |

Table VIII: Description of the variables used in the probit and multiple linear regression models. All variables used were dummy variables except for time, 'round' and interaction terms with round

| Predictor | Description |
|---|---|
| Above | =1 if question asks for temperature *above* X deg. C, 0 if it asks for *below* |
| Above & certain | =1 if question asked for temperature above X deg. C & just the point forecast is shown (format A) |
| Above & uncertainty | =1 if question asked for temperature above X deg. C & subject is shown uncertainty information (formats B &C) |
| Area | 1= questions where the greatest area between the high/low range and the asked temperature does not get the correct answer for Group C participants for questions 1,12 and 13 |
| Business | =1 for participants studying business/economics related subjects and 0 if science/humanities |
| Checks internet for weather forecast | Whether or not participant checks weather forecasts on the internet (1=checks internet for weather, 0= checks other sources) |
| Checks weather at least every 2-3 days | 1=if participant checks weather at least every2-3 days, 0= checks weather weekly or never |
| Die question mistake | 1= participant answered the question about rolling a fair die twice incorrectly |
| Early day correct | =1 if statement A was the most probable outcome ( e.g. Monday as opposed to Wednesday) |
| Early day correct & Format A | =1 if statement A was the most probable outcome and participants were shown just a point forecast without uncertainty information |
| Early day correct & Format C | =1 if statement A was the most probable outcome and participants were shown a graph with uncertainty information |
| Easy question | 1= question is easy, 0=swing/ hard |
| English is first language | Whether or not English is the first language (1= English is first language, 0=otherwise) |
| Format A | 1= participant was presented with a table without uncertainty information |
| Format B | 1= participant was presented with a table with uncertainty information |
| Format C | 1= participant was presented with a bar graph with uncertainty information |
| Format B with Description | =1 for group B participants with test questions phrased as '*most likely*' & were given tables mentioning that the maximum temperature ranges fall within the indicated range 9 out of 10 times, 0 for all other participants |
| Hard question | 1= question is hard, 0=swing/easy |
| Humanities | =1 for participants humanities |
| Length | 1= questions where the greatest length between the high/low range & the asked temperature does not get the correct answer for Group B participants for questions 1, 12 and 13. |
| Male | Gender of participant (1=male, 0= female ) |
| Order 1 | Equals one if question order is: 1, 2,..., 20 |

| | |
|---|---|
| Order 1 & Format B | =1 if question order is 1 & participants were shown Format B |
| Order 1 & Format C | =1 if question order is 1 & participants were shown Format C |
| Order 2 | Equals one if question order is: 20, 19,..., 1 |
| Order 2 & Format B | =1 if question order is 2 & participants were shown Format B |
| Order 2 & Format C | =1 if question order is 2 & participants were shown Format C |
| Order 4 | Equals one if question order is: 10, 9,..., 1, 20, 19,..., 11 |
| Order 4 & Format B | =1 if question order is 4 & participants were shown Format B |
| Order 4 & Format C | =1 if question order is 4 & participants were shown Format C |
| Round number | Round number |
| Round number & Format B | Round number * Format B |
| Round number & Format C | Round number * Format C |
| Round number squared | Round number squared |
| Round number squared & Format B | Round number squared * Format B |
| Round number squared & Format C | Round number squared * Format C |
| Sample question mistake | 1= if participant answered two of the test questions incorrectly |
| Science | =1 for participants studying science/ engineering |
| Swing question | 1=question is a swing question, |
| Swing question & Format A | =1 if question was swing & participants were shown just a point forecast without uncertainty information |
| Swing question & Format C | =1 if question was swing &participants were shown a graph with uncertainty information |
| Test question dummy | =1 if test questions were phrased as '*most likely*' & 0 if phrased as '*expected*' for Format B and C participants |
| Time | average response time taken by a participant to respond for each of the 20 'lotteries |

Figure 7: Average response time for each of the groups by order and question type. Group B participants who were provided with a table with uncertainty information took more time despite the question type compared to Group A and Group C participants except when the question order is 4, where Group C participants take more time on average.



## References


Anderson, J.C. and Mueller, J.M., 2005. The Effects Of Experience And Data Presentation Format On An Auditing Judgment. Journal of Applied Business Research (JABR), 21(1): 53-63.

Back, W.E., Boles, W.W. and Fry, G.T., 1999. Defining triangular probability distributions from historical cost data. Journal of Construction Engineering and Management, 126(1): 29-37.

Benbasat, I. and Dexter, A.S., 1985. An Experimental Evaluation of Graphical and Color-Enhanced Information Presentation. Management Science, 31(11): 1348-1364.

Cardinaels, E., 2008. The interplay between cost accounting knowledge and presentation formats in cost-based decision-making. Accounting, Organizations and Society, 33(6): 582-602.

Chua, H., Yates, J. and Shah, P., 2006. Risk avoidance: Graphs versus numbers. Memory & Cognition, 34(2): 399-410.

Dickson, G.W., DeSanctis, G. and McBride, D.J., 1986. Understanding the effectiveness of computer graphics for decision support: a cumulative experimental approach. Communications of the ACM, 29(1): 40-47.

Driver, R. et al., 2010. Helping Consumers Understand Investment Risk. Asssociation of British Insurers (ABI).

Evans, M., Hastings, N. and Peacock, B., 2000. Statistical Distributions. Wiley-Interscience, New York

Fleming, N. and Mills, C., 1992. Not Another Inventory, Rather a Catalyst for Reflection. To Improve the Academy, 11: 137.

Ghani, E.K., Laswad, F., Tooley, S. and Jusoff, K., 2009. The Role of Presentation Format on Decision-makers' Behaviour in Accounting. International Business Research 2(1).

Hesse, R., 2000. In The Classroom. Triangle Distribution: Mathematica Link for Excel. Decision Line 31(3).

Johnson, D., 1997. The triangular distribution as a proxy for the beta distribution in risk analysis. Journal of the Royal Statistical Society: Series D (The Statistician), 46(3): 387-398.

Kotz, S., 2004. Beyond Beta: Other Continuous Families of Distributions with Bounded Support and Applications. World Scientific Publishing. Co. Pte. Ltd Singapore

Larham, R., 2010. Project Costing with the Triangular Distribution and Moment Matching, Math Help Forum Journal

Lusk, E.J. and Kersnick, M., 1979. The Effect of Cognitive Style and Report Format on Task Performance: The MIS Design Consequences. Management Science, 25(8): 787-798.

NRC, 2006. Completing the Forecast: Characterizing and Communicating Uncertainty for Better Decisions Using Weather and Climate Forecasts National Academies Press, Wachington D.C.

Remus, W., 1984. An Empirical Investigation of the Impact of Graphical and Tabular Data Presentations on Decision Making. Management Science, 30(5): 533-542.

Roulston, M.S., Bolton, G.E., Kleit, A.N. and Sears-Collins, A.L., 2006. A Laboratory Study of the Benefits of Including Uncertainty Information in Weather Forecasts. Weather and Forecasting, 21: 116-122.

Roulston, M.S. and Kaplan, T.R., 2009. A laboratory-based study of understanding of uncertainty in 5-day site-specific temperature forecasts. Meteorological Applications, 16(2): 237-244.

So, S. and Smith, M., 2003. The impact of presentation format and individual differences on the communication of information for management decision making. Managerial Auditing Journal, 18 (1): 59 - 67.

Speier, C., 2006. The influence of information presentation formats on complex task decision-making performance. International Journal of Human-Computer Studies, 64(11): 1115-1131.
Sullivan, J.J., 1988. Financial Presentation Format and Managerial Decision Making. Management Communication Quarterly, 2(2): 194-216.
Vessey, I. and Galletta, D., 1991. Cognitive Fit: An Empirical Study of Information Acquisition. Information Systems Research, 2(1): 63-84.
Weisstein, E.W., Triangular Distribution, MathWorld-A Wolfram Web Resource.
Winett, R.A. and Kagel, J.H., 1984. Effects of Information Presentation Format on Resource Use in Field Studies. The Journal of Consumer Research, 11(2): 655-667.
Zmud, R.W., 1978. An Empirical Investigation of the Dimensionality of the Concept of Information. Decision Science, 9(2): 187-195.