Supplementary Materials for Tannenbaum, Fox, & Ülkümen "Judgment Extremity and Accuracy under Epistemic versus Aleatory Uncertainty"

Contents

1 Study 2S: Differences in Evidence Sensitivity Across Domains

In Study 2S we sought to complement the findings in Study 2 by focusing on a smaller number of domains but more exhaustively sampling items within each domain. We selected three judgment domains expected to span the range of perceived epistemicness — geography questions, weather estimates, and upcoming NBA basketball games. Based on the findings of Study 2, we expected that questions about U.S. geography would be viewed as primarily epistemic, NBA basketball games would be viewed as primarily aleatory, and weather-related events would be viewed between these extremes.

Study 2S Methods

We recruited a sample of 37 self-identified basketball fans[1] from MTurk (19% female, mean age = 33 years, range: 19–59 years) who participated in return for a small cash payment plus entry into a drawing to receive an NBA basketball jersey of their choice. One participant reported using outside sources while completing the task, and was excluded from the analysis.

First, participants provided probability judgments to 16 two-alternative questions for each of three domains: (i) outcomes of upcoming NBA playoff games, (ii) outcomes of upcoming daytime high temperatures, and (iii) general knowledge questions about the relative geographic sizes of different U.S. states (see Table 1 for sample questions). As before, one alternative was designated as focal for each question (counterbalanced across participants) and participants estimated the probability from 0-100% that the focal team would win their game (basketball), the focal city would register a higher daytime high temperature (weather), or the focal state is geographically larger (geography). For each domain we constructed two sets of four targets (e.g., Atlanta, Buffalo, Los Angeles, Memphis; Miami, Minneapolis, New Orleans, San Francisco) and defined events by factorially pairing one target from each set (e.g., Atlanta-Miami) so that we had sixteen possible pairings for each domain. The ordering of judgment domains and questions within domains

[1] We planned to sample 50 basketball fans, but successfully recruited only 37 participants before the start of the NBA playoffs. Note that the design of Study 2S was entirely within-subject, so the sample still provided reasonable statistical power.

Table 1: Study 2B sample questions

Domain	Question
Basketball	Suppose that the Los Angeles Clippers play the Boston Celtics in the NBA finals. What do you think is the probability that the Los Angeles Clippers will win?
Temperature	Consider a day picked at random next year in Los Angeles and Minneapolis. What do you think is the probability that it will be warmer in Los Angeles that day?
Geography	Consider the geographical size of Wisconsin and Georgia. What do you think is the probability that Wisconsin is the larger state?

Table 2: Epistemicness Ratings and Judgment Extremity in Study 2S

		Judgment Extremity			
	Epistemicness M (SD)	MAD from $p = .50$	median $p > .50$	median $p < .50$	proportion $p = 0$ or 1
Geography	6.19 (1.03)	.28	.99	.10	.36
Temperature	3.89 (1.22)	.23	.80	.25	.14
Basketball	3.23 (1.14)	.20	.70	.30	.10

were randomized, with the constraint that all judgments within a domain were completed before advancing to the next block. Afterwards, participants were provided with a list of the targets from each domain and were asked to assess their relative strength (strength of teams, warmth of cities, size of states) following the same protocol as in Study 2. In the final phase of the study participants rated each domain for its degree of epistemicness. A single trial was selected at random from each domain and participants rated the event using a 10-item EARS scale similar to that in Study 1 (Cronbach's α ranged from .83 to .88 across domains).

Study 2S Results

Table 2 provides average epistemicness ratings for the three domains. Our NBA basketball fans rated basketball as the least epistemic domain, followed by temperature estimates, and then by geography questions — all consistent with the pattern that we had observed for sports versus weather versus geography in Study 2. All means were reliably different from one another ($p < .001$ for all pairwise comparisons).

More importantly, and consistent with the results of Study 2, both judgment extremity and evidence sensitivity followed the same rank-ordering as epistemicness ratings — smallest for basketball, intermediate for city temperature, and highest for the state geography questions. For judgment extremity, mean absolute deviations were .20 for basketball, .23 for weather, and .28 for geography questions. All domains reliably differed from one another ($p \leq .001$ for all pairwise comparisons). We find a similar pattern across the three judgment domains when restricting the analysis to judgments above .50, below .50, or when dichotomizing responses into certain versus uncertain responding (p-values $\leq .015$ for all pairwise comparisons).

We also recovered estimates of evidence sensitivity for each domain using the specification outlined in eq. (7), and then performed pairwise comparisons between domains. As expected,[2] evidence sensitivity was smallest for basketball predictions ($k = 1.54$), intermediate for temperature estimates ($k = 2.05$), and largest for geography questions ($k = 2.81$). All estimates were significantly different from one another (p-values $< .005$).

[2]We observe a similar rank-ordering across domains when analyzing evidence sensitivity over participants and over items, rather than over trials as we did in the foregoing analysis.

Finally, we examined the relationship between rated epistemicness and evidence sensitivity. At the trial-level our hypothesis would predict a positive interaction between strength ratings and perceived epistemicness — the slope on strength ratings, which represents an estimate of evidence sensitivity, should increase as perceived epistemicness increases. We therefore regressed judged log odds onto log strength ratios, epistemicness ratings, and the interaction between the two. As expected, we found a reliable and positive interaction effect ($b_{intx} = 0.36$, SE = 0.04, $p < .001$). Based on the regression coefficients, estimates of evidence sensitivity would be expected to increase from 1.51 to 2.73 when going from one standard deviation below to one standard deviation above the mean in rated epistemicness.

2 Knowledge and Sensitivity to Evidence Strength: Additional Analyses

In Study 1, in which we did not elicit strength ratings for NCAA basketball teams, we can nevertheless examine the interaction between rated knowledge and judged epistemicness on judgment extremity. The foregoing arguments suggest that we should observe greater sensitivity of judged probabilities to differences in team strength (i.e., steeper slopes in Figure 1) for participants who consider themselves more knowledgeable. To do this we took the average ratings from our initial three screening measures as our proxy for self-reported knowledge. We then regressed absolute differences of judged probabilities from $1/2$ (i.e., degree of judgment extremity) onto these knowledge scores, participants' epistemicness ratings, and the interaction between the two, with game fixed effects and participant random effects. Even though the range on our knowledge ratings was restricted to the higher end of the scale (since we screened participants on these items), we nevertheless observe a positive interaction effect, $b = .02$, SE = .009, $p = .015$. That is, epistemicness was a stronger predictor of judgment extremity for more knowledgeable participants (e.g., 1 standard deviation above the mean; $b = .054$, SE = .013, $p < .001$) than it was for less knowledgeable participants (i.e., 1 standard deviation below the mean; $b = .019$, SE = .01, $p = .105$).

For Study 2, in which we elicit a single probability, pair of strength ratings, and knowledge rating for each of 6 domains (out of 12) per participant, we only have power to perform a coarse analysis that averages knowledge and epistemicness ratings over all domains. Nevertheless, this analysis also reveals a pattern consistent with the present analysis: the rank-order correlation between k and epistemicness monotonically increases from the lowest to highest quartile of mean knowledge ratings (Spearman's ρ was .03, .14, .17, and .22, respectively). Alternatively, using an analytic measure of evidence sensitivity (by dividing judged log odds by log strength ratios) we can estimate this relationship at the trial level by regressing analytic k values onto epistemicness and knowledge ratings. Because analytic k values vary widely, we used robust regression in which estimates are robust to outliers or influential observations. Using this alternative approach we again find a postive interaction effect, $b = .02$, SE = .01, $p = .051$. Perceived epistemicness was a stronger predictor of evidence sensitivity when participants felt knowledge about the task domain (e.g., 1 standard deviation above the mean; $b = .29$, SE = .05, $p < .001$) than it was when participants felt less knowledgable about the task domain (i.e., 1 standard deviation below the mean; $b = .18$, SE = .05, $p < .001$).

3 NBA Study Materials

Welcome, and thank you for your participation! This is a research study on how people think about upcoming basketball games.

In this survey you will provide estimates for three NBA basketball games. Afterward we will ask you some additional questions. This study should take 3 minutes or less to complete.

For the questions that follow, it is important that you do not consult outside sources (such as EPSN, Wikipedia, etc). Just give the best answer off the top of your head.

Sample Judgment from Singular Frame:

The Chicago Bulls will play the Detroit Pistons on March 21st.

What is the probability that the Bulls will win?

Sample Judgment from Distributional Frame:

The Chicago Bulls will play the Detroit Pistons on February 20th, March 21st, and April 3rd.

What is the probability that the Bulls will win on March 21st?

Epistemicness Ratings

Consider a typical NBA basketball game. Which team will win that game is ... (1 = *not at all*, 7 = *very much*)

1. ... something that has an element of randomness.
2. ... knowable in advance, given enough information.
3. ... determined by chance factors.
4. ... something that well-informed people would agree on.

Strength Ratings

Listed below are all six NBA teams you responded to earlier in this survey.

First, choose the team you believe is the strongest of the six, and set that team's strength rating to 100. Assign the remaining team ratings in proportion to the strength of the strongest team. For

example, if you believe that a given team is half as strong as the strongest team (the team you gave 100), give that team a strength rating of 50.

- Chicago Bulls
- Detroit Pistons
- Memphis Grizzlies
- LA Clippers
- Toronto Raptors
- Charlotte Hornets

All other questions

1. On average, how many NBA games per week have you watched or listened to this season?
2. Approximately how many NBA games have you watched or listened to so far this entire season?
3. On average, how many hours per week do you spend reading, watching, or listening to sports news/commentary about NBA basketball?
4. Which NBA teams do you root for? (check all that apply)
5. What is your age?
6. What is your gender?

***

As a way of saying thank you for taking this survey, you will be entered into a raffle to receive an NBA jersey of your choice.

If you would like to be considered for the raffle, please leave your email address below. You will only be contacted if you win the raffle. (We will delete all email addresses from our servers once the raffle is done)

***

The survey will be complete after clicking the continue button below.

Feel free to let us know in the space below if you have any comments or suggestions for the researchers who designed this survey.

4 Study 1 Materials

Initial Screening Questions

1. To what extent do you follow the NCAA men's basketball this season? (1 = *not at all*, 7 = *very much so*)
2. To what extent do you consider yourself a fan of NCAA men's basketball? (1 = *I don't consider myself a fan at all*, 7 = *I am an avid fan*)
3. How knowledgeable do you feel about NCAA men's basketball this season? (1 = *not at all knowledgeable*, 7 = *extremely knowledgeable*)

Instructions

Welcome. For this study, you will make predictions for the 2015 NCAA Men's Basketball tournament. This entire study should take about 10 to 15 minutes to complete.

***

HERE IS THE BASIC SET-UP:

For each game you will provide a probability estimate between 0 and 100.

- For example, a 30% probability means that you think a given team would win their game 3 times out of 10; a 70% probability means that you think that the team would win their game 7 times out of 10; and so on.

Please provide your best judgment, as accuracy will be rewarded.

- In addition to the flat study payment, some participants will be selected at random to receive a bonus payment of up to $100. The amount of the bonus will depend on the accuracy of their judgments. We will make these payments once the study is complete and all results can be calculated.
- How is accuracy scored? We do this by calculating the mean squared difference between each participant's probability judgments and the actual outcomes (in this case, which team won each game). This is a common measure for measuring accuracy, as you do best when you express your true beliefs and are as accurate as possible.
- Thus, it is in your best interests to express your true beliefs and be as accurate as possible.

Once you have read the instructions above and are ready to begin, click the continue button below.

Sample Judgment Item

Game: 1 of 28

In the first round of the NCAA tournament **Wisonsin** will play **Coastal Carolina**. Wisonsin is the 1st seed and Coastal Carolina is the 16th seed in the West region.

What is the probability that **Wisconsin** will win?

Epistemicness Ratings

Earlier you were asked the following question: [randomly sampled question from previous set of trials]

Determining which team will win ... (1 = *Not at all*, 7 = *Very much*)

1. ... is in principle knowable in advance
2. ... is something that has an element of randomness
3. ... is something that has been determined in advance
4. ... is unpredictable
5. ... is knowable in advance, given enough information
6. ... is determined by chance factors
7. ... could play out in different ways on similar occasions
8. ... is something that well-informed people would agree on
9. ... is something that could be better predicted by consulting an expert
10. ... is something that becomes more predictable with additional knowledge or skills

All other questions

1. On average, how many NCAA basketball games per week have you watched or listened to this season?
2. Approximately how many NCAA basketball games have you watched or listened to so far this entire season?
3. On average, how many hours per week do you spend reading, watching, or listening to sports news/commentary about NCAA basketball?

4. Which NCAA basketball team do you primarily root for?
5. What is your age?
6. What is your gender?
7. Did you use external sources, such as ESPN.com or Wikipedia, while completing any of the questions during this study? (Please be honest. You will be paid regardless of your response)
8. Do you have any last comments or suggestions you would like to share with the researchers who designed this study? Please let us know what you think.

5 Study 2 Materials

General Instructions

- In what follows, you will be presented with a series of questions.
- Your task is to estimate the probability that various events will occur or are true. Please do your best to be as accurate as possible. For example, if you say that the probability is 40% then you should be correct 4 times out of 10; likewise if you say that the probability is 60% then you should be correct 6 out of 10 times, and so on.
- In this study we are interested in your own personal judgment. Therefore it is important that you rely only on your own knowledge and give your best answer "off the top of your head." Please do NOT rely on any external sources (e.g., books, web pages, other people).

Sample Judgment Item

Consider the weather in Chicago and Minneapolis.

What is the probability that there will be more rainy days next May in Chicago than Minneapolis?

Strength Elicitation Instructions

Rain
Consider Chicago and Minneapolis. First, choose the city you believe is the **rainier** of the two cities, and set that city's rainy rating to 100. Assign the other city a rainy rating in proportion to the first city. For example, if you believe that a given city is half as rainy as the first city (the one you gave 100), give that city a rainy rating of 50.

Temperature
Consider Pittsburgh and Portland. First, choose the city you believe is the **warmer** of the two cities, and set that city's warmth rating to 100. Assign the other city a warmness rating in proportion to the first city. For example, if you believe that a given city is half as warm as the first city (the one you gave 100), give that city a warmth rating of 50.

Politics
Consider Barack Obama and Mitt Romney. First, choose the presidential candidate you believe is **a stronger candidate in terms of his prospects to win the 2012 election**, and set that candidate's strength rating to 100. Assign the other candidate a strength rating in proportion to the first

candidate. For example, if you believe that a given candidate is half as strong as the first candidate (the one you gave 100), give that candidate a strength rating of 50.

Football

Consider the Arizona Cardinals and the San Francisco 49ers. First, choose the football team you believe is the **stronger** of the two teams, and set that team's strength rating to 100. Assign the other team a strength rating in proportion to the first team. For example, if you believe that a given team is half as strong as the first team (the one you gave 100), give that team a strength rating of 50.

Baseball

Consider the Chicago Cubs and the LA Dodgers. First, choose the baseball team you believe is the **stronger** of the two teams, and set that team's strength rating to 100. Assign the other team a strength rating in proportion to the first team. For example, if you believe that a given team is half as strong as the first team (the one you gave 100), give that team a strength rating of 50.

Movies

Consider the two upcoming movies *The Amazing Spider-Man* and *The Dark Night Rises*. First, choose the movie you believe is **more highly anticipated** of the two movies, and set that movie's strength rating to 100. Assign the other movie a strength rating in proportion to the first movie. For example, if you believe that a given movie is half as highly anticipated as the first move (the one you gave 100), give that movie a strength rating of 50.

Housing

Consider Atlanta and Nashville. First, choose the city you believe is the **more expensive (in terms of housing prices)** of the two cities, and set that city's expensiveness rating to 100. Assign the other city an expensiveness rating in proportion to the first city. For example, if you believe that a given city is half as expensive as the first city (the one you gave 100), give that city an expensiveness rating of 50.

Crime

Consider Columbus and Detroit. First, choose the city you believe is the **more violent** of the two cities, and set that city's violence rating to 100. Assign the other city a violence rating in proportion to the first city. For example, if you believe that a given city is half as violent as the first city (the one you gave 100), give that city a violence rating of 50.

Geography

Consider Nevada and Wyoming. First, choose the state you believe is the **larger (in sq. miles)** of the two states, and set that state's size rating to 100. Assign the other state a size rating in proportion to the first state. For example, if you believe that a given state is half as large as the first

state (the one you gave 100), give that state a size rating of 50.

Populations
Consider Istanbul, Turkey and Shanghai, China. First, choose the city you believe is the **more populous** of the two cities (i.e. the one that has the higher population), and set that city's size rating to 100. Assign the other city a size rating in proportion to the first city. For example, if you believe that a given city is half as populous as the first city (the one you gave 100), give that city a size rating of 50.

Soccer
Consider the Italian national soccer team and the German national soccer team. First, choose the soccer team you believe is the **stronger** of the two teams, and set that team's strength rating to 100. Assign the other team a strength rating in proportion to the first team. For example, if you believe that a given team is half as strong as the first team (the one you gave 100), give that team a strength rating of 50.

Ocean Size
Consider the Atlantic Ocean and the Indian Ocean. First, choose the ocean you believe is the **larger (in sq. miles)** of the two oceans, and set that ocean's size rating to 100. Assign the other ocean a size rating in proportion to the first ocean. For example, if you believe that a given ocean is half as large as the first ocean (the one you gave 100), give that ocean a size rating of 50.

Epistemicness Ratings

Recall the following question: [randomly presented question from previous set of trials]

Please rate this question along the following dimensions: (1 = *Not at all*, 7 = *Very much*)

1. The answer to this question is in principle knowable in advance
2. The answer to this question has an element of randomness
3. The answer to this question has been determined in advance
4. The answer to this question feels unpredictable

Knowledge Ratings

Recall the following question: [randomly presented question from previous set of trials]

How knowledgable do you feel when answering this question? (0 = *not at all knowledgable*, 10 = *extremely knowledgable*)

Direction Check

Note: This was included in our survey but we did not filter responses according to the direction check.

SPORTS PARTICIPATION

Individual preferences and knowledge, along with situational variables, can greatly impact the decision process. In order to facilitate our research on decision making we are interested in knowing certain factors about you, the decision maker. Specifically, we are interested in whether you actually take the time to read the directions; if not, then the data we collect based on your responses will be invalid. So, in order to demonstrate that you have read the instructions, please ignore the sports items below. Instead, simply click on the title at the top of this screen (i.e., "sports participation"). The title should turn green if you have clicked on it correctly. Thank you very much.

Which of these activities do you engage in regularly? (click on all that apply)

- skiing
- football
- soccer
- swimming
- snowboarding
- tennis
- running
- basketball
- hockey
- cycling

All other questions

1. What is your age?
2. What is your gender?
3. Did you use external sources, such as Wikipedia, when answering any of the questions during this study? (Please be honest. You will be paid regardless of your response)
4. Finally, are there any general comments or questions that you have about the study?

6 Study 2S Materials

General Instructions

In what follows, you will be presented with a series of uncertain events or statements.

- Your task is to estimate the probability that various events will occur or facts are true. Please do your best to be as accurate as possible. For example, if you say that the probability is 40% then you should be correct 4 times out of 10; likewise if you say that the probability is 60% then you should be correct 6 out of 10 times, and so on.
- In this study we are interested in your own personal judgment. Therefore it is important that you rely only on our own knowledge and give your best answer "off the top of your head." Please do NOT rely on any external sources (e.g., books, web pages, other people).

Sample Judgment Items

Basketball
Suppose that the **San Antonio Spurs** play the **Philadelphia 76ers** in the NBA finals.

What do you think is the probability that the **San Antonio Spurs** will win?

Temperature
Consider a day picked at random next year in **Minneapolis** and **Buffalo**.

What do you think is the probability that it will be warmer in **Minneapolis** that day?

Geography
Consider the geographical size of **Georgia** and **Wisconsin**.

What do you think is the probability that **Georgia** is the larger state?

Strength Elicitation Instructions

Basketball
Consider the eight basketball teams remaining in the NBA playoffs. First, choose the team you believe is the **strongest** of the eight, and set that team's strength to 100. Assign the remaining teams ratings in proportion to the strength of the strongest team. For example, if you believe that a given team is half as strong as the strongest team (the team you gave 100), give that team a strength rating of 50.

- Boston Celtics
- Indiana Pacers
- Los Angeles Clippers
- Los Angeles Lakers
- Miami Heat
- Oklahoma City Thunder
- Philadelphia 76ers
- San Antonio Spurs

Temperature

Consider the following eight US cities below. First, choose the city you believe is the **warmest** (year-round average) of the eight, and set that city's warmth rating to 100. Assign the remaining cities ratings in proportion to the warmth of the warmest city. For example, if you believe that a given city is half as warm as the warmest city (the city you gave 100), give that city a warmth rating of 50.

- Atlanta
- Buffalo
- Los Angeles
- Miami
- Memphis
- Minneapolis
- New Orleans
- San Francisco

Geography

Consider the following eight US states below. First, choose the state you believe is the **largest** (i.e., greatest geographical area) of the eight, and set that state's size rating to 100. Assign the remaining states ratings in proportion to the size of the largest state. For example, if you believe that a given state is half as large as the largest state (the state you gave 100), give that state a size rating of 50.

- Georgia
- Idaho
- Kansas

- Maryland
- Massachusetts
- Montana
- New Mexico
- Wisconsin

Knowledge Ratings

1. Please rate your knowledge of **US city climates** (temperature) from 0 (completely ignorant) to 100 (extremely expert).
2. Please rate your knowledge of **US state sizes** (geographical size) from 0 (completely ignorant) to 100 (extremely expert)
3. Please rate your knowledge of **NBA basketball** from 0 (completely ignorant) to 100 (extremely expert).
4. On average, how much time do you spend reading about or watching NBA basketball (ESPN, Sports Illustrated, sports blogs, etc.) on a weekly basis?
5. About how many NBA basketball games have you watched or listened to so far this season?

Epistemicness Ratings

Please read the question below and rate it along the following dimensions: [randomly sample question from previous set of trials]

1. The outcome to this question is in principle knowable in advance
2. Determining the outcome to this question depends on knowledge or skill
3. With enough information, one could know the answer to this question in advance
4. The outcome of this question feels unpredictable
5. The outcome of this question has an element of randomness
6. The outcome of this question feels like it is determined by chance factors
7. The outcome of this question could play out in different ways on similar occasions
8. Well-informed people would agree on what the outcome to this question would be
9. The outcome to this question has been determined in advance

10. If I could consult an expert on this topic it would improve my prediction

Direction Check

Note: This was included in our survey but we did not filter responses according to the direction check.

Please read the directions below:

Most modern theories of decision making recognize the fact that decisions do not take place in a vacuum. Individual preferences and knowledge, along with situational variables can greatly impact the decision process. In order to facilitate our research on decision making we are interested in knowing certain factors about you, the decision maker. Specifically, we are interested in whether you actually take the time to read the directions; if not, then the data we collect based on your responses will be invalid. So, in order to demonstrate that you have read the instructions, please ignore the preferences form below, and simply write "I read the instructions" in the box below. Thank you very much.

Which of these activities do you engage in regularly? (click on all that apply)

- organized sports
- watching TV
- listening to music
- gardening
- crafts
- photography
- surfing the web
- playing video games
- working out
- reading for fun
- yoga/meditation
- cooking

All other questions

1. How old are you?
2. What is your the primary ethnicity that you identify with?

3. For any of the basketball games, did you take home-court advantage into consideration when making your probability judgment?
4. If you answered Yes to the question above, indicate how often you took this information into consideration.
5. Did you use external sources, such as Wikipedia, when answering any of the questions during this study?
6. Finally, are there any general comments, questions, or suggestions that you have about the study?

7 Study 3 Materials

General Instructions

- For this study, you will be presented with pairs of US cities and estimate the probability that one city has a higher temperature than the other city.
- Please do your best to be as accurate as possible. For example, if you say that the probability is 40% then you should be correct 4 times out of 10; likewise if you say that the probability is 60% then you should be correct 6 out of 10 times, and so on.
- Please do not use external sources, such as Google or Wikipedia, to answer these questions. Just give us the best answer "off the top of your head."

Sample Judgment Items

Historic Average format
What is the probability that the the average temperature last year was higher in **Indianapolis** than in **Anchorage**?

Arbitrary Day format
What is the probability that the temperature of an arbitrarily-selected day from last year was higher in **Indianapolis** than in **Anchorage**?

Strength Elicitation Instructions

Consider the following six US cities:

- Anchorage
- Indianapolis
- Minneapolis
- Phoenix
- San Diego
- San Francisco

First, choose the city you believe is the warmest (year-round average) of the six, and set that city's warmth rating to 100. Assign the remaining cities ratings in proportion to the warmth of the warmest city. For example, if you believe that a given city is half as warm as the warmest city (the city you gave 100), give that city a warmth rating of 50.

Epistemicness Ratings

Earlier you were asked the following question: [randomly sampled question from previous set of trials]

Determining the answer to this question ... (1 = *Not at all*, 7 = *Very much*)

1. ... is in principle knowable in advance
2. ... is something that has an element of randomness
3. ... is something that has been determined in advance
4. ... feels unpredictable
5. ... is knowable in advance, given enough information
6. ... feels like it is determined by chance factors
7. ... feels like it could play out in different ways on similar occasions
8. ... is something that well-informed people would agree on
9. ... is something that could be better predicted by consulting an expert
10. ... is something that becomes more predictable with additional knowledge or skills

Knowledge Ratings

For this study you were repeatedly asked to judge which of two US cities had a higher temperature. How knowledgeable did you feel as you were performing this task? (1 = *not at all knowledgable*, 10 = *extremely knowledgable*)

All other questions

1. What is your age?
2. What is your gender?
3. Did you use external sources, such as Google or Wikipedia, when answering any of the questions during this study? (Please be honest. You will be paid regardless of your response)
4. Finally, are there any general comments or questions that you have about the study?

8 Study 4 Materials

Task Instructions/Prime

Pattern Detection Instructions

The first part of this study is as follows:

- On each trial, you will try to predict which of two events, X or O, will occur next.
- The sequence of Xs and Os has been set in advance, and your task is to figure out this pattern.

*** [page break] ***

- Prior to each trial, you will predict whether an X or and O will appear next. After each prediction, you will view the actual outcome on the screen.
- There will be a total of 168 trials. In addition to the standard payment for participation, the participant who identifies the most outcomes correctly will receive an addition bonus of $25. THIS IS A REAL CASH BONUS. The $25 will be paid out once the study is complete.
- Also keep in mind that the proportion of Xs or Os may not necessarily be equal — there may be more relatively more Xs or relatively more Os.

*** [page break] ***

- To get yourself familiar with the task, you will now perform 10 practice trials. For each trial, choose either X or O.
- You will learn the actual outcome immediately after making your prediction.

*** [after practice trial] ***

- Very good. Now you will advance to the real trials.
- There are a total of 168 trials. You will complete these trials in 2 blocks (84 trials apiece), with a short rest in between.
- Press the continue button when you are ready ...

Random Guessing Instructions

The first part of this study is as follows:

- On each trial, you will try to guess which of two events, X or O, will occur next.
- The order of Xs and Os will be randomly generated by a computer program, and your task is to guess which outcome will appear next.

*** [page break] ***

- Prior to each trial, you will guess whether an X or and O will appear next. After each guess, you will view the actual outcome on the screen.
- There will be a total of 168 trials. In addition to the standard payment for participation, the participant who guesses the most outcomes correctly will receive an addition bonus of $25. THIS IS A REAL CASH BONUS. The $25 will be paid out once the study is complete.
- Also keep in mind that the proportion of Xs or Os may not necessarily be equal — there may be more relatively more Xs or relatively more Os.

*** [page break] ***

- To get yourself familiar with the task, you will now perform 10 practice trials. For each trial, choose either X or O.
- You will learn the actual outcome immediately after making your guess.

*** [after practice trial] ***

- Very good. Now you will advance to the real trials.
- There are a total of 168 trials. You will complete these trials in 2 blocks (84 trials apiece), with a short rest in between.
- Press the continue button when you are ready ...

Instructions for Judgment Task

- You have now finished the first part of this study. Well done!
- For the second part of this study, you will be presented with a series of questions.
- Your task is to estimate the probability that various events will occur. Please do your best to be as accurate as possible.
- For example, if you say that the probability is 40% then you should be correct 4 times out of 10; likewise if you say that the probability is 60% then you should be correct 6 out of 10 times, and so on.

Sample Judgment Item

Consider the daily high temperature this July 1st in **Austin, TX** and **San Francisco, CA**.

What is the probability that it will be hotter in **Austin, TX**?

Strength Elicitation Instructions

Consider the following eight cities below. First, choose the city you believe is the **warmest** (year-round average) of the eight, and set that city's warmth rating to 100. Assign the remaining cities ratings in proportion to the warmth of the warmest city. For example, if you believe that a given city is half as warm as the warmest city (the city you gave 100), give that city a warmth rating of 50.

- New York, NY
- San Diego, CA
- San Francisco, CA
- Austin, TX
- Albuquerque, NM
- Colorado Springs, CO
- Tulsa, OK
- Wichita, KS

Epistemicness Ratings

Earlier you were asked the following question: [randomly sampled question from previous set of trials]

Determining which city will be warmer ... (1 = *Not at all*, 7 = *Very much*)

1. ... is in principle knowable in advance
2. ... is something that has an element of randomness
3. ... is something that has been determined in advance
4. ... feels unpredictable
5. ... is knowable in advance, given enough information
6. ... feels like it is determined by chance factors
7. ... feels like it could play out in different ways on similar occasions
8. ... is something that well-informed people would agree on
9. ... is something that could be better predicted by consulting an expert
10. ... is something that becomes more predictable with additional knowledge or skills

Knowledge Ratings

Consider the task you just engaged in, where you made a series of temperature predictions between two US cities. In general, how knowledgable did you feel as you were performing this task? (1 = *not at all knowledgable*, 10 = *extremely knowledgable*)

Now, please rate your knowledge for each of the following eight cities. (1 = *not at all knowledgable*, 10 = *extremely knowledgable*)

- New York, NY
- San Diego, CA
- San Francisco, CA
- Austin, TX
- Albuquerque, NM
- Colorado Springs, CO
- Tulsa, OK
- Wichita, KS

All other questions

1. What is your age?
2. What is your gender?
3. Is English your first language?

9 Testing for Binary Complementarity

One implication of support theory is *binary complementarity*, namely that for all binary partitions a hypothesis and its complement should sum to unity:

$$p(A,B) + p(B,A) = \frac{s(A)}{s(A)+s(B)} + \frac{s(B)}{s(B)+s(A)} = 1 \tag{1}$$

In all studies participants were asked to assign probabilities to either p(A,B) or p(B,A) for each trial. We were therefore able to perform a between-subject test of binary complementarity for each study. We compared $p(A,B)$ and $1-p(B,A)$ for each of question item per study using two-sample t-tests (see also Fox, 1999). Tables 3–9 below report the results for Studies 1–4, respectively. Each table displays mean judged probabilities for $p(A,B)$ and $1-p(B,A)$, along with test statistics comparing the two sets of judgments.

Table 3: Study 1 Test of Binary Complementarity

Team A	Team B	$p(A,B)$	$1-p(B,A)$	t-statistic	p-value
Villanova	Lafayette	.83	.78	−1.18	.24
Virginia	Belmont	.78	.82	0.98	.33
Oklahoma	Albany	.73	.73	−0.04	.97
Louisville	UC Irvine	.74	.73	−0.24	.81
Northern Iowa	Wyoming	.63	.64	0.35	.73
Michigan State	Georgia	.63	.57	−2.05	.04
North Carolina State	LSU	.53	.46	−2.51	.01
Kansas	New Mexico State	.76	.77	0.46	.65
Notre Dame	Northeastern	.78	.72	−1.55	.12
Maryland	Valparaiso	.73	.67	−1.74	.08
West Virginia	Buffalo	.71	.67	−1.45	.15
Butler	Texas	.55	.55	−0.08	.94
Wichita State	Indiana	.58	.55	−1.03	.30
Cincinnati	Purdue	.54	.50	−1.38	.17
Gonzaga	North Dakota State	.78	.77	−0.22	.82
Iowa State	UAB	.71	.71	0.00	1.00
Georgetown	Eastern Washington	.69	.67	−0.52	.61
Utah	Stephen F. Austin	.65	.67	0.58	.56
SMU	UCLA	.56	.56	0.17	.87
Iowa	Davidson	.59	.58	−0.47	.64
San Diego State	St. John's	.53	.50	−1.11	.27
Wisconsin	Coastal Carolina	.85	.76	−2.19	.03
Arizona	Texas Southern	.82	.81	−0.22	.82
Baylor	Georgia State	.70	.70	0.01	.99
North Carolina	Harvard	.72	.76	1.08	.28
Arkansas	Wofford	.68	.71	1.04	.30
VCU	Ohio State	.50	.48	−0.54	.59
Oregon	Oklahoma State	.55	.47	−3.13	< 0.01

Table 4: Study 2 Test of Binary Complementarity

Domain	Target A	Target B	$p(A,B)$	$1-p(B,A)$	t-statistic	p-value
Baseball	Cubs	Dodgers	.44	.47	−0.76	.45
Box Office	Spider Man	Dark Knight	.35	.39	−0.72	.47
Crime Rates	Detroit	Colombus	.79	.77	0.29	.78
Football	49ers	Cardinals	.54	.59	−1.5	.14
Housing Prices	Nashville	Atlanta	.44	.46	−0.21	.83
Ocean Size	Atlantic	Indian	.80	.79	0.20	.84
City Population	Istanbul	Shanghai	.18	.22	−0.83	.41
Rainfall	Chicago	Minneapolis	.50	.55	−1.13	.26
Soccer	Italy	Germany	.52	.49	0.79	.43
Geography	Nevada	Wyoming	.42	.53	−1.52	.13
Temperature	Portland	Pittsburgh	.37	.35	0.34	.73
Politics	Obama	Romney	.56	.61	−1.14	.26

Table 5: Study 2S Test of Binary Complementarity (Basketball)

Target A	Target B	$p(A, B)$	$1 - p(B, A)$	t-statistic	p-value
76ers	Clippers	.48	.39	1.09	.28
76ers	Lakers	.24	.20	0.67	.51
76ers	Spurs	.37	.42	−0.44	.66
76ers	Thunder	.37	.48	−0.95	.35
Celtics	Clippers	.62	.70	−1.23	.23
Celtics	Lakers	.44	.43	0.19	.85
Celtics	Spurs	.56	.62	−0.78	.44
Celtics	Thunder	.59	.53	0.71	.48
Heat	Clippers	.55	.64	−1.12	.27
Heat	Lakers	.47	.42	0.85	.40
Heat	Spurs	.64	.60	0.71	.48
Heat	Thunder	.64	.68	−0.51	.61
Pacers	Clippers	.44	.37	1.10	.28
Pacers	Lakers	.23	.30	−1.19	.24
Pacers	Spurs	.43	.39	0.58	.57
Pacers	Thunder	.46	.41	0.50	.62

Table 6: Study 2S Test of Binary Complementarity (Weather)

Target A	Target B	$p(A, B)$	$1 - p(B, A)$	t-statistic	p-value
Atlanta	Miami	.32	.34	−0.26	.79
Atlanta	Minneapolis	.85	.80	1.12	.27
Atlanta	New Orleans	.43	.43	−0.04	.97
Atlanta	San Francisco	.70	.41	3.40	$< .01$
Buffalo	Miami	.16	.10	0.85	.40
Buffalo	Minneapolis	.47	.59	−1.55	.13
Buffalo	New Orleans	.12	.18	−1.26	.22
Buffalo	San Francisco	.28	.21	1.02	.31
Los Angeles	Miami	.46	.44	0.23	.82
Los Angeles	Minneapolis	.85	.90	−1.00	.33
Los Angeles	New Orleans	.67	.40	3.28	$< .01$
Los Angeles	San Francisco	.65	.64	0.07	.94
Memphis	Miami	.34	.22	2.26	.03
Memphis	Minneapolis	.81	.81	0.01	.99
Memphis	New Orleans	.43	.42	0.11	.92
Memphis	San Francisco	.54	.55	−0.11	.91

Table 7: Study 2S Test of Binary Complementarity (Geography)

Target A	Target B	$p(A, B)$	$1 - p(B, A)$	t-statistic	p-value
Georgia	Kansas	.52	.47	0.51	.61
Georgia	Maryland	.82	.85	−0.43	.67
Georgia	New Mexico	.31	.30	0.06	.96
Georgia	Wisconsin	.49	.37	1.18	.25
Idaho	Kansas	.50	.45	0.55	.59
Idaho	Maryland	.79	.63	1.46	.16
Idaho	New Mexico	.37	.22	1.90	.07
Idaho	Wisconsin	.39	.47	−0.70	.49
Massachusetts	Kansas	.20	.18	0.21	.83
Massachusetts	Maryland	.56	.39	1.57	.13
Massachusetts	New Mexico	.27	.06	2.61	.01
Massachusetts	Wisconsin	.15	.22	−0.64	.53
Montana	Kansas	.69	.72	−0.39	.70
Montana	Maryland	.85	.96	−1.36	.19
Montana	New Mexico	.66	.63	0.29	.77
Montana	Wisconsin	.74	.76	−0.19	.85

Table 8: Study 3 Test of Binary Complementarity

City A	City B	$p(A, B)$	$1 - p(B, A)$	t-statistic	p-value
Indianapolis	Anchorage	.66	.83	−7.64	$< .01$
Anchorage	Minneapolis	.28	.41	−5.43	$< .01$
Phoenix	Anchorage	.89	.94	−3.05	$< .01$
San Diego	Anchorage	.82	.90	−4.14	$< .01$
San Francisco	Anchorage	.75	.87	−5.25	$< .01$
Indianapolis	Minneapolis	.51	.64	−6.68	$< .01$
Indianapolis	Phoenix	.17	.19	−1.29	0.20
Indianapolis	San Diego	.30	.35	−1.93	0.05
San Francisco	Indianapolis	.58	.65	−3.13	$< .01$
Phoenix	Minneapolis	.83	.88	−3.04	$< .01$
Minneapolis	San Diego	.25	.31	−2.47	0.01
Minneapolis	San Francisco	.31	.36	−2.36	0.02
Phoenix	San Diego	.69	.69	0.03	0.98
Phoenix	San Francisco	.75	.74	0.69	0.49
San Diego	San Francisco	.57	.63	−2.87	$< .01$

Table 9: Study 4 Test of Binary Complementarity

Target A	Target B	$p(A, B)$	$1 - p(B, A)$	t-statistic	p-value
Albuquerque	Austin	.54	.46	1.77	.08
CO Springs	Albuquerque	.38	.45	−1.47	.15
Albuquerque	New York	.63	.69	−1.09	.28
Albuquerque	San Diego	.60	.57	0.66	.51
Albuquerque	San Francisco	.73	.72	0.12	.91
Tulsa	Albuquerque	.38	.43	−1.17	.25
Wichita	Albuquerque	.40	.44	−0.94	.35
Austin	CO Springs	.64	.66	−0.56	.58
New York	Austin	.32	.26	1.24	.22
San Diego	Austin	.36	.37	−0.35	.73
San Francisco	Austin	.28	.27	0.28	.78
Austin	Tulsa	.63	.67	−0.96	.34
Wichita	Austin	.41	.36	1.15	.25
CO Springs	New York	.54	.60	−1.40	.16
San Diego	CO Springs	.50	.61	−2.31	.02
CO Springs	San Francisco	.55	.60	−0.93	.35
CO Springs	Tulsa	.44	.46	−0.56	.58
CO Springs	Wichita	.47	.50	−0.85	.40
San Diego	New York	.58	.66	−1.68	.10
San Francisco	New York	.47	.51	−0.87	.39
Tulsa	New York	.57	.61	−1.04	.30
Wichita	New York	.58	.64	−1.43	.16
San Diego	San Francisco	.70	.75	−1.34	.18
Tulsa	San Diego	.48	.44	0.76	.45
San Diego	Wichita	.52	.57	−0.93	.36
Tulsa	San Francisco	.58	.65	−1.39	.17
San Francisco	Wichita	.34	.37	−0.63	.53
Tulsa	Wichita	.49	.52	−1.02	.31