Version: November 20, 2018

SUPPLEMENTARY MATERIALS
Intentionally "Biased": People Purposely Use To-Be-Ignored Information, But Can Be Persuaded Not To

Data, code for all analyses, original study materials, and preregistrations are posted on this OSF page: https://osf.io/uvqhs/

Contents of Supplementary Materials

Supplement 1: Breakdown of participants and exclusions

Number who…	Study 1a	Study 1b	Study 1c[1]	Study 2	Study 3	Study 4a	Study 4b
Started Survey	442	521	462	194	602	763	903
Failed Attention Check	24	41	40	9	85	43	69
Completed At Least One Comprehension Check Question[2]	N/A	442	409	163	N/A	628	732
Failed Comprehension Check	N/A	118	131	37	N/A	127	140
Completed DV	393	295	274	113	511	480	571
Completed Survey	393	292	273	113	501	480	567

[1] We made a mistake in qualtrics when closing incomplete survey responses after Study 1c had completed. We accidentally deleted 50 incomplete responses instead of "closing' them. All reported results exclude those 50 responses that we (unfortunately) cannot recover.
[2] Studies 1a and 3 did not include comprehension checks.

Supplement 2: Study 1a detailed write-up

Procedures. Participants who passed the attention check first read the historical passage from Fischhoff (1975) describing a conflict between the British and the Gurka in Nepal:

> "For some years after the arrival of Hastings as governor-general of India, the consolidation of British power involved serious war. The first of these wars took place on the northern frontier of Bengal where the British were faced by the plundering raids of the Gurkas of Nepal. Attempts had been made to stop the raids by an exchange of lands, but the Gurkas would not give up their claims to country under British control, and Hastings decided to deal with them once and for all. The campaign began in November 1814. It was not glorious. The Gurkas were only some 12,000 strong; but they were brave fighters, fighting in territory well-suited to their raiding tactics. The older British commanders were used to war in the plains where the enemy ran away from a resolute attack. In the mountains of Nepal it was not easy even to find the enemy. The troops and transport animals suffered from the extremes of heat and cold, and the officers learned caution only after sharp reverses. Major-General Sir D Octerlony was the one commander to escape from these minor defeats." (p. 298)[3]

On the next page all participants learned "additional information" – the Gurka won the conflict. Participants were randomly assigned to either read that a group of MTurk workers or university students had answered a multiple choice question about the outcome of the conflict. The question had the options that Fischhoff (1975) used: a) The Gurka won, b) The British won, c) Military stalemate with no peace settlement, d) Military stalemate with a peace settlement. Next, participants completed the dependent measure, which asked "If your task were to estimate the shares of the [university students/other MTurk workers] who gave an answer of 'a', 'b', 'c', and 'd' to the multiple choice question, do you think that using the 'additional information' shown above (The Gurka won this conflict) would make you more or less accurate?" Their options were: "I think my answer would be more accurate if I <u>used the 'additional information'</u>

[3] We copied this passage from Fischhoff (1975). Fischhoff (1975) took the passage from Woodward's (1938) The Age of Reform (p.383-384).

to some degree" and "I think my answer would be more accurate if I completely ignored the 'additional information'." Participants also reported their age, sex, and level of education.

Results. We found that 73% (287/393) of participants thought that their answer would be more accurate if they used the outcome information. Participants' answers were not meaningfully different when they were making estimates of MTurk workers' responses (74%, 146/198) and university students' responses (72%, 141/195), $\chi 2(1, N = 393) = 0.10$, $p = .750$.

Supplement 3: Study 1b detailed write-up

Procedures. Participants who passed the attention check were informed that the survey had a number of comprehension checks, that they would earn $0.10 for reading the instructions and completing at least one comprehension check, and that they would earn an additional $0.25 for passing all comprehension checks and completing the survey. On the next page, participants learned that a university student, John, had predicted the proportion of other university students giving each answer to a four option multiple choice question about history. Participants read that John had made two estimates and we wanted to know which they thought was more accurate.

Next, participants took the first comprehension check. It consisted of two multiple choice questions that asked what John's occupation was (university student) and how many estimates he made (2). Participants who passed the comprehension check learned that John read the historical passage from Fischhoff (1975) describing a conflict between the British and the Gurka in Nepal and read the passage themselves. On the next page, participants completed the second comprehension check. It consisted of two multiple choice questions that asked who the Gurkas were in conflict with (the British) and where the conflict was (Nepal).

On the next page, participants learned about John's first estimate. John indicated what percent of students he thought chose each of the answers to a multiple choice question about the outcome of the conflict without learning the actual outcome. The question was the same one used in Study 1a. On the next page, participants learned about John's second estimate. For this estimate, John learned the actual outcome of the conflict. Participants were randomly assigned to learn that the actual outcome was either: The Gurka won, The British won, Military stalemate with no peace settlement, or Military stalemate with a peace settlement. Then John once again indicated the percent of students he thought chose each outcome.

Finally, participants completed the dependent measure. Participants were informed that John gave different answers for his first and second estimate, and reported whether they thought his first or second estimate was more accurate. Their options were "I think that estimate #1 (John's estimate before he was told the outcome of the conflict) was more accurate" and "I think that estimate #2 (John's estimate after he was told the outcome of the conflict) was more accurate." Participants also reported their age, sex, and level of education.

Results. We found that 67% (198/295) of participants thought that John's estimate was more accurate after he learned the outcome information. Participants' responses did not significantly differ based on the "actual" outcome of the conflict: Gurka victory (73%, 54/74), British victory (61%, 45/74), Stalemate (73%, 54/74), Stalemate with a peace settlement (62%, 45/73), $\chi 2(3, N = 295) = 4.62$, $p = .201$.

Supplement 4: Study 1c detailed write-up

Procedures. Study 1c was similar to Study 1b with a few differences. First, participants read that their job was to report which of two university students, John and Robert, made more accurate estimates of other university students' responses to a multiple choice question. We randomized which name appeared first in the instructions, and orthogonally randomized which name was associated with learning the outcome information. Second, participants learned that one of the university students made his estimate after learning the outcome information and the other student made his estimate without learning the outcome information. Third, the dependent measure asked participants who they thought was more accurate in predicting other university students' answers. The options were "I think that John [Robert] (who was **not** told the outcome of the conflict) was more accurate" and "I think that Robert [John] (who was told the outcome of the conflict) was more accurate."

Results. We found that 62% (171/274) of participants thought that the participant who learned the outcome information was more accurate. Participants' responses did not significantly differ based on the "actual" outcome of the conflict: Gurka victory (58%, 40/69), British victory (65%, 44/68), Stalemate (59%, 41/69), Stalemate with a peace settlement (68%, 46/68), $\chi 2(3, N = 295) = 1.79$, $p = .617$. These results are consistent with Study 1b, in which participants reported that a person would make a more accurate estimate after (vs. before) learning the outcome of the conflict.

Supplement 5: Study 2 detailed procedures

Procedures. Participants who passed the attention check were informed that the survey had a number of comprehension checks, that they would earn $0.20 for reading the instructions and completing at least one comprehension check, and that they would earn an additional $0.30 for passing all comprehension checks and completing the survey. On the next page, participants read that they would learn about estimates that two finance students ("Participant A" and "Participant B") made during an exercise, and tell us which student they think made a better estimate. We orthogonally manipulated whether "Participant A" or "Participant B" appeared first and which one would later receive to-be-disregarded information.

Next, participants took the first comprehension check. It consisted of two multiple choice questions that asked how many people's estimates they would answer questions about (2) and what Participant A and Participant B study (finance). Participants who passed the comprehension check learned about the estimates. In Stage 1, a group of finance students were told the earnings that a company had had for the last 10 years, and then they predicted what that company would earn the following quarter. In Stage 2, several months later, we asked two new students, Participant A and Participant B, to make estimates about what their fellow students had predicted earlier in Stage 1. On the next page, participants completed the second comprehension check. It consisted of two multiple choice questions that asked what the large group of finance students predicted (a company's earnings) and what Participant A and Participant B estimated (the large group of finance students' predictions about the company's earnings).

On the next page, participants read that the key difference between Participant A and Participant B is that one learned the company's actual earnings and the other did not. Finally, participants completed the dependent measure. Participants were informed that Participant A and

Participant B gave different answers, and reported who they thought was more accurate in estimating the large group of finance students estimates. Their options were “I think that Participant A [B] (who was **not** told the company’s earnings) was more accurate” and “I think that Participant B [A] (who was told the company’s earnings) was more accurate.” Participants also reported their age, sex, and level of education.

Supplement 6: Study 3 detailed procedures

Procedures. Participants who passed the attention check read that they would answer 10 trivia questions, learn about their performance, and then make a prediction. Participants learned that they would have 20 seconds to answer each question before the survey auto-advanced, and were asked not to look up the answers to the trivia questions. Next, participants completed 10 trivia questions that were randomly selected out of a pool of 11 (Supplement 7 for a full list of the questions).

After completing all 10 questions, participants learned how many questions they answered correctly. Participants also read that they would be matched with a past participant, learn how many of the questions the past participants answered correctly, and estimate how the past participants answered an 11th trivia question. On the next page participants learned that the past participant got the same number of correct answers that they did for the first 10 questions.

On the next page, participants saw the 11th trivia question that the past participant had answered. The 11th question was always the question that the participant had not answered previously. At the bottom of this page, all participants answered two yes/no questions that solicited their plan to use the to-be-disregarded information that half of participants were randomly assigned to receive. These questions read “Do you think that you could more accurately predict the past participant's answer if you learned the correct answer to this question?” and “If you learned the correct answer to this question, would you use that information to predict the past participant's answer?” On the next page, the half of participants who were randomly assigned to learn the correct answer read the 11th question again and then learned the correct answer.

Finally, participants answered the study's dependent measure. We asked participants how they thought the past participant answered the question. Participants were provided with four empty boxes next to the four answers to the question. In each box, they indicated the chances (reported in percentages) that the past participant had chosen that answer. They could not advance unless all of their responses summed to 100%. On the next page participants answered three exploratory questions. The first asked whether they thought the past participant or themselves is better at answering trivia questions on a five-point Likert scale. The second asked participants why they planned to use or ignore the to-be-disregarded information with an open ended text box. The third question was a yes/no question that asked whether or not participants had looked up answer to any questions. Participants also reported their age, sex, and level of education.

Supplement 7: Trivia questions used in Study 3

We thank Michael O'Donnell for sharing many of these questions with us.

Which artist painted 'Guernica,' depicting scenes from the Spanish Civil War?
- Dali
- Picasso
- Matisse
- Cezanne

Who was the fourth president of the United States?
- James Monroe
- Thomas Jefferson
- James Madison
- John Quincy Adams

Where is the Great Victoria Desert located?
- Canada
- West Africa
- Australia
- India

The ulna is a long bone in which part of the body?
- Arm
- Foot
- Leg
- Neck

What element is common to all acids?
- Hydrogen
- Carbon
- Sulphur
- Oxygen

What is the largest country in the world?
- Canada
- Russia
- USA
- China

Who is the father of Geometry?
- Aristotle
- Euclid
- Pythagoras
- Kepler

What was the first battle of the Civil War?
- The Battle of Aquia Creek
- The First Battle of Bull Run
- The Battle of Fort Sumter
- The Battle of Antietam

Which leader of the Soviet Union resigned on December 25, 1991?
- Boris Yeltsin
- Nikita Khrushchev
- Alexander Popov
- Mikhail Gorbachev

The Hindu festival Diwali is known as the festival of what?
- Harvests
- Lights
- Rains
- Winds

What was the English term for a German submarine in World War II?
- G-Boat
- U-Boat
- S-Boat
- H-Boat

Supplement 8: Additional analyses for Study 3

Preregistered DV

3) Dependent variable. Describe the key dependent variable(s) specifying how they will be measured.
Forecasters will give probabilities that the target forecaster will answer each of four possible answers: A,B,C,D.
We will measure hindsight bias as a function of the predicted (p%) and the actual share (s%) of respondents answering correctly, in particular:

DV1: p
DV2: (p-s)/s

DV1 is our favored dependent variable, but we will report analyses for both.

Preregistered analyses

5) Analyses. Specify exactly which analyses you will conduct to examine the main question/hypothesis.
Participants will answer two questions (before completing the DV) that ask: "Do you think that you could more accurately predict the past participant's answer if you learned the correct answer to this question?" and "If you learned the correct answer to this question, would you use that information to predict the past participant's answer?" Our analyses will be conditional on the answers to these questions.

Main Analysis
T-test (separately for DV1 and DV2) excluding participants who answered "yes" to 1) either of the questions above, 2) only the first question, and 3) only the second question (i.e. 3 separate analyses).

6) More analyses. Any secondary analyses?
Regressions (run separately for DV1 and DV2) testing for an interaction between an indicator (0/1) of whether or not the participant learned the correct answer and an indicator (0/1) of whether or not the participant answered "yes" to 1) either of the questions in Section 5, 2) only the first question, and 3) only the second question (i.e. 3 separate analyses).

T-test (separately for DV1 and DV2) only including participants who reported that they are 1) equally as good as, and 2) as good as or better than the respondent who they were matched with at answering trivia questions (2 separate analyses).

Analyses

Main pre-registered analysis - T-tests excluding participants who answered "**yes**" to…

Answered "yes" to…	**DV**	**Result**
Either question	DV1	t(181) = 0.17, p = .867
Question 1	DV1	t(277) = -0.95, p = .342
Question 2	DV1	t(209) = -0.84, p = .404
Either question	DV2	t(181) = -0.94, p = .350
Question 1	DV2	t(277) = -1.33, p = .184
Question 2	DV2	t(209) = -1.56, p = .119

Additional analysis - T-tests excluding participants who answered "**no**" to…

Answered "no" to…	**DV**	**Result**
Both questions	DV1	t(326) = -5.23, p < .001
Question 1	DV1	t(230) = -4.99, p < .001
Question 2	DV1	t(298) = -4.62, p < .001
Both questions	DV2	t(326) = -4.14, p < .001
Question 1	DV2	t(230) = -4.53, p < .001
Question 2	DV2	t(298) = -3.70, p < .001

Secondary pre-registered analysis – Regressions testing for interaction

Answered "yes" to…	**DV**	**Result**
Either question	DV1	t(507) = 3.28, p = .001
Question 1	DV1	t(507) = 3.21, p = .001
Question 2	DV1	t(507) = 2.36, p = .019
Either question	DV2	t(507) = 1.84, p = .066
Question 1	DV2	t(507) = 2.36, p = .019
Question 2	DV2	t(507) = 1.44, p = .152

Secondary pre-registered analysis – T-tests based on perceived relative performance

Including participants who thought that they were…	**DV**	**Result**
Equally as good as past participant	DV1	$t(360) = -2.92$, $p = .003$
As good as or better than past participant	DV1	$t(424) = -4.11$, $p < .001$
As good as or worse than past participant	DV1	$t(445) = -2.97$, $p = .003$
Equally as good as past participant	DV2	$t(360) = -2.64$, $p = .009$
As good as or better than past participant	DV2	$t(424) = -3.55$, $p < .001$
As good as or worse than past participant	DV2	$t(445) = -3.13$, $p = .002$

Supplement 9: Study 4a detailed procedures

Procedures. Participants who passed the attention check were informed that the survey had a number of comprehension checks, that they would earn $0.15 for reading the instructions and completing at least one comprehension check, and that they would earn an additional $0.25 for passing all comprehension checks and completing the survey. On the next page, participants learned that they would read about a criminal trial in which the defendant has been charged with murdering two people, and asked their opinion about aspects of the trial as if they were a juror.

Next, participants took the first comprehension check. It consisted of two multiple choice questions that asked how many people's decisions they would answer questions about (2) and what the charge was in the trial (murder). Participants who passed the comprehension check learned more about the trial. Participants read that the case involves a man charged with murdering his estranged wife and male neighbor. The prosecutor had claimed that the defendant killed the victims in a fit of jealous rage. The defendant has said he had found the bodies when he returned to his former home to retrieve personal papers. On the next page, participants completed the second comprehension check. It consisted of two multiple choice questions that asked how the defendant was related to the first victim (estranged wife) and how the defendant was related to the second victim (neighbor). Next, all participants read a police officer's testimony and read one of three statements that either did or did not discredit the evidence depending on their condition:

> Up to this point, the evidence has been incomplete and ambiguous.
>
> Then a police officer's testimony is presented. It consists of the officer reading off a transcript made from a wiretapped telephone conversation in which the defendant appears to confess to a friend minutes after fleeing the scene ("I killed Mary and some bastard she was with. God, I don't . . . yeah, I ditched the blade").
>
> At this point the defense lawyer objects.
>
> [Admissible condition]

> The judge overrules the objection and admits the evidence as valid.
>
> [Inadmissible/due-process condition]
> The judge sustains the objection and admonishes the jury to disregard the evidence because the phone was intercepted without a warrant, deeming it an illegal search.
>
> [Inadmissible/unreliable condition]
> The judge sustains the objection and admonishes the jury to disregard the evidence because the original audio is of poor quality (it is barely understandable) and hence it is difficult to determine what was actually said during the call. The transcript that was read is not trustworthy.

Next, participants answered two questions, the dependent measure and the mediator, in randomized order. The dependent measure asked participants if they would vote "guilty" or "not guilty" when submitting your verdict for the court. Their options were “I would vote ‘not guilty’ and I would vote ‘guilty’”. The mediator asked whether they planned to use or disregard the wiretap evidence. Their options were “My plan is [was] to completely disregard the police officer’s testimony when deciding on my verdict” and “My plan is/was to take the police officer’s testimony into account when deciding on my verdict”. Finally, participants reported their age, sex, and level of education.

Supplement 10: Study 4b detailed procedures

Procedures. Study 4b was similar to Study 4a with a few changes. First, we only asked for guilty verdicts and did not ask about participants' plan to ignore or attend to the inadmissible evidence. Second, we added a control condition that submitted a guilty verdict without learning about the wiretap evidence. Third, we added two additional Due-process conditions (Strong Due-process, Very Strong Due-process) that made stronger due process arguments than the standard Due-process condition:

> [Strong Due-process condition]
>
> The judge sustains the objection and admonishes the jury to disregard the evidence because the phone was intercepted without a warrant, deeming it an illegal search.
>
> The judge explains that it is very important not to consider evidence that was gathered with an illegal search. In fact, illegal searches violate constitutional rights and threaten your own individual privacy.
>
> [Very Strong Due-process condition]
>
> The judge sustains the objection and admonishes the jury to disregard the evidence because the phone was intercepted without a warrant, deeming it an illegal search.
>
> The judge explains that it is very important not to consider evidence that was gathered with an illegal search. In fact, illegal searches violate constitutional rights and threaten your own individual privacy. If we allowed such evidence, the police will conduct searches proactively, routinely violating our rights and probably especially those of minorities. For example, the police could pull you over for a routine traffic stop and search your glove compartment, phone, trunk, etc. without probable cause, and anything they found could be used against you.

Supplement 11: Study S4a

Participants. We ran this study on Amazon Mechanical Turk. Participants were required to pass an attention check in order to participate in the study, earned $0.15 for starting the study and completing at least one of the study's two comprehension checks, and earned an additional $0.25 for passing the study's comprehension checks and completing the study. Overall, 796 participants clicked on the survey link, 61 participants failed the attention check, 626 participants completed at least one of the study's comprehension checks, 106 participants failed a comprehension check, 504 participants passed both comprehension checks and completed the dependent measure, and 502 participants submitted completed survey responses. The final sample averaged 33 years of age and was 53% male.

Procedures. Participants who passed the attention check were informed that the survey had a number of comprehension checks, that they would earn $0.15 for reading the instructions and completing at least one comprehension check, and that they would earn an additional $0.25 for passing all comprehension checks and completing the survey. On the next page, participants read that we would ask them which of two people on a jury, Juror #1 and Juror #2, would make a better decision regarding the guilt of a defendant charged with murder in a criminal trial. We orthogonally manipulated whether "Juror #1" or "Juror #2" appeared first and which one would later use to-be-disregarded information.

Next, participants took the first comprehension check. It consisted of two multiple choice questions that asked how many people's decisions they would answer questions about (2) and what the charge was in the trial (murder). Participants who passed the comprehension check learned more about the trial. Participants read that the case involves a man charged with murdering his estranged wife and male neighbor. The prosecutor had claimed that the defendant

killed the victims in a fit of jealous rage. The defendant has said he had found the bodies when he returned to his former home to retrieve personal papers. On the next page, participants completed the second comprehension check. It consisted of two multiple choice questions that asked how the defendant was related to the victim (estranged wife) and how many victims there were (2). Next, all participants read about the evidence and one of three statements that either did or did not discredit the evidence:

> Up to this point, the evidence has been incomplete and ambiguous.
>
> Then a police officer's testimony is presented. It consists of the officer reading off a transcript made from a wiretapped telephone conversation in which the defendant appears to confess to a friend minutes after fleeing the scene ("I killed Mary and some bastard she was with. God, I don't . . . yeah, I ditched the blade").
>
> [Admissible condition]
> The judge overrules the objection and admits the evidence as valid.
>
> [Inadmissible/due-process condition]
> The judge sustains the objection and admonishes the jury to disregard the evidence because the phone was intercepted without a warrant, deeming it an illegal search.
>
> [Inadmissible/unreliable condition]
> The judge sustains the objection and admonishes the jury to disregard the evidence because the original audio is of poor quality (it is barely understandable) and hence it is difficult to determine what was actually said during the call. The transcript that was read is not trustworthy.

Next, participants read that one of the jurors decided to use the wiretap evidence from the wiretap when deciding on their verdict, and the other decided to completely ignore the wiretap evidence. Then participants reported which juror they thought would make a better decision regarding the defendant's guilt. Their options were "Juror #1 [#2], who chose to **use** the police officer's testimony (i.e. the wiretap), will make a better decision" and "Juror #2 [#1], who chose to completely **ignore** the police officer's testimony (i.e. the wiretap), will make a better decision". Finally, participants reported their age, sex, and level of education.

Results. We found that participants' preferences for either using or ignoring the wiretap evidence followed the same general pattern as the guilty verdicts from Kassin & Sommers (1997). Kassin & Sommers (1997) found that participants in the Admissible, Due-process, and

Unreliable conditions voted guilty 79%, 55%, an 24% of the time respectively, and we found that participants in our Admissible, Due-process, and Unreliable conditions expressed a preference for using the wiretap evidence 94%, 62%, an 33% of the time respectively, $\chi 2$ (1, N=504) = 134.59, $p < .001$. These percentages are significantly different in all between condition comparisons, $\chi 2$ (1, N>=335) >= 28.04, $p < .001$. In study 4a in the paper, we measure both guilty verdicts and intention to attend to the inadmissible evidence in the same study. We find that participants' plan to attend to, or ignore, the inadmissible evidence mediates their decision to give a guilty or not guilty verdict.

Study S4b

Overview. In study S4b, we investigated whether or not experimenter demand drove the effectiveness of the Very Strong due-process discrediting message from Study 4b. Because this discrediting message contained a strong plea, it is possible that participants voted "not guilty" more often because they believed that it was the outcome that the experimenter wanted. To test this possibility, we ran a similar study to Study 4b with two pieces of evidence that participants read about - 1) the same piece of incriminating evidence from Study 4b and 2) a much more trivial piece of evidence. If the success of the Very Strong Due-process condition was due to experimenter demand, we would expect that 1) participants would show a strong reaction to this message even if only the more trivial evidence was discredited and 2) discrediting both the incriminating and more trivial evidence with this message might have a similar effect. In contrast to these possibilities, we found that 1) discrediting only the trivial piece of evidence with the Very Strong Due-process message had virtually no effect and 2) discrediting the incriminating evidence with the Strong Due-process massage had a stronger effect than discrediting the more trivial evidence.

Participants. We ran this study on Amazon Mechanical Turk. Participants were required to pass an attention check in order to participate in the study, earned $0.15 for starting the study and completing at least one of the study's two comprehension checks, and earned an additional $0.25 for passing the study's comprehension checks and completing the study. Overall, 519 participants clicked on the survey link, 15 participants failed the attention check, 437 participants completed at least one of the study's comprehension checks, 90 participants failed a comprehension check, 321 participants passed both comprehension checks and completed the

dependent measure, and 319 participants submitted completed survey responses. The final sample averaged 37 years of age and was 48% male.

Procedures. Study S4b was similar to Study 4b with a few key differences. First, participants learned that the witness read transcripts from two separate wiretapped phone conversations. The "incriminating" conversation was the same as the message from Study 4b and the "trivial" conversation was a new addition.

> [incriminating tape]
>
> Tape #1 is made from a wiretapped telephone conversation in which the defendant appears to confess to a friend minutes after fleeing the scene ("I killed Mary and some bastard she was with. God, I don't . . . yeah, I ditched the blade").
>
> At this point the defense lawyer objects on the grounds that the phone call was intercepted without a warrant.
>
> [trivial tape]
>
> Tape #1 is made from a wiretapped telephone conversation in which the defendant speaks to a friend about his feelings for his estranged wife three months before the crime ("I am really mad at Mary, I can't believe she left me after all these years").
>
> At this point the defense lawyer objects on the grounds that the phone call was intercepted without a warrant.

We counterbalanced the order in which the two conversations were presented. We also orthogonally manipulated whether each conversation was discredited with the Very Strong Due-process message or ruled to be admissible by the judge immediately after the conversation was presented (see below). This resulted in a 2 (Incriminating conversation first, Trivial conversation first tape second & trivial first) x 4 (Trivial conversation discredited, Incriminating conversation discredited, Neither conversation discredited, Both conversations discredited) design.

> [Very Strong Due-process]
>
> The judge sustains the objection and admonishes the jury to disregard the evidence in tape #1 because this phone call was intercepted without a warrant, deeming it an illegal search.
>
> The judge explains that it is very important not to consider evidence that was gathered with an illegal search. In fact, illegal searches violate constitutional rights and threaten your own individual privacy. If we allowed such evidence, the police will conduct searches proactively, routinely violating our rights and probably especially those of minorities. For example, the police

> could pull you over for a routine traffic stop and search your glove compartment, phone, trunk, etc. without probable cause, and anything they found could be used against you.
>
> [Admissible]
>
> The judge overrules the objection because the police did have a warrant for the wiretap of the phone conversation in tape #1. He provides the defense lawyer with a copy of the warrant and admits the evidence.

Results. Participants had very different reactions to the incriminating and trivial evidence being discredited with the Very Strong Due-process message, which is consistent with the notion that the Very Strong Due-process message did not cause participants to vote "not guilty" because of its strength alone. Participants who saw only the trivial conversation discredited with the strong message voted guilty at virtually the same (very high) rate as participants who saw neither conversation discredited (94% vs. 95%), $\chi 2$ (1, N=161) = 0.10, p = .746. In contrast, participants who saw only the incriminating conversation discredited voted guilty at a much lower rate than participants who saw neither conversation discredited (67% vs. 95%), $\chi 2$ (1, N=161) = 20.78, p < .001.

We ran a logistic regression of participants' verdict on two dummies indicating whether 1) the trivial conversation was discredited and 2) the incriminating conversation was discredited. We ran a test of the hypothesis that the coefficients of the two dummies are equal, and found that the coefficient of the "incriminating evidence discredited" dummy was significantly larger than that of the "trivial evidence discredited" dummy, $\chi 2$ (1, N=321) = 16.92, p < .001. Given this evidence, we find it unlikely that experimenter demand to vote not guilty drove the success of the Very Strong Due-process message in Study 4b.