Research Article

Psychological Science
1–10

DOI: 10.1177/0956797619856844
www.psychologicalscience.org/PS

The Illusion of Consensus: A Failure to Distinguish Between True and False Consensus

TC

Sami R. Yousif, Rosie Aboody, and Frank C. Keil
Department of Psychology, Yale University

Abstract

When evaluating information, we cannot always rely on what has been presented as truth: Different sources might disagree with each other, and sometimes there may be no underlying truth. Accordingly, we must use other cues to evaluate information—perhaps the most salient of which is consensus. But what counts as consensus? Do we attend only to surface-level indications of consensus, or do we also probe deeper and consider why sources agree? Four experiments demonstrated that individuals evaluate consensus only superficially: Participants were equally confident in conclusions drawn from a true consensus (derived from independent primary sources) and a false consensus (derived from only one primary source). This phenomenon was robust, occurring even immediately after participants explicitly stated that a true consensus was more believable than a false consensus. This illusion of consensus reveals a powerful means by which misinformation may spread.

Keywords

consensus, conformity, social learning, reasoning, open data, open materials, preregistered

Received 9/10/18; Revision accepted 5/7/19

Imagine that you are trying to decide what you think about a newly proposed social policy. You ask several of your coworkers for their thoughts, and each one mentions their support for the policy. Although you were not sure what to think before, you have now heard the same information repeated several times: There appears to be a consensus that the policy is a positive change. But now imagine that you discover that your colleagues received their information from the same source (i.e., a single coworker) instead of reaching their conclusions independently. Would you be less likely to take their stance seriously?

In theory, at least, consensus is a strong cue to a claim's trustworthiness: If many sources all report the same story, it is reasonable to assume it is likely to be true. Indeed, a large literature has demonstrated that adults, children, and diverse species of animals are highly sensitive to consensus (e.g., Asch, 1956; Corriveau & Harris, 2010; Kundey et al., 2012). But is every consensus equally informative? Although in the above example there was clearly a consensus about the policy (in that multiple sources appeared to support it), it does not seem to be a very good consensus. Indeed, this consensus is illusory, as there is ultimately only a single source of information. Do individuals distinguish between different types of consensus, or do they take any form of agreement at face value?

The Consequences of Consensus

Classic work on conformity suggests that people may overrely on apparent consensus—even when that consensus is obviously wrong (Asch, 1956)—and expect consensus to align with their own beliefs (Ross, Greene, & House, 1977). This sensitivity to consensus emerges early in life, with children as young as 3 years old often

Corresponding Authors:
Sami Yousif, Yale University, Department of Psychology, Box 208205, New Haven, CT 06520-8205
E-mail: sami.yousif@yale.edu

Rosie Aboody, Yale University, Department of Psychology, Box 208205, New Haven, CT 06520-8205
E-mail: rosie.aboody@yale.edu

aligning themselves with the majority when there is disagreement (Corriveau, Fusaro, & Harris, 2009; Corriveau & Harris, 2010; Fusaro & Harris, 2008). More recent work has investigated how and under what conditions children are susceptible to consensus (Burdett et al., 2016; Hu, Buchsbaum, Griffiths, & Xu, 2013; Hu, Whalen, Buchsbaum, Griffiths, & Xu, 2015). Yet little work has examined sensitivity to the quality of consensus, even though this seems directly relevant to a wide range of psychological and sociological questions (but see Einav, 2014, 2018). Consensus is critical not only as we decide which rumors to take seriously but also as we interpret academic articles, news sources, and virtually any kind of information for which there can be a consensus at all. Therefore, to understand how people come to believe (and defend) erroneous information, we must first understand how they reason about consensus.

Outside the realm of cognitive science, one study has shed light on a timely example of false consensus (and its consequences). This study focused on the sociology of climate-change denial in order to better understand the gap between scientific consensus and public opinion. Notably, more than 80% of the studied climate-change-denying blogs relied on a single primary source (an individual who, despite having never conducted any relevant research, claims to be an expert on polar bears; Harvey et al., 2018). The assertions of that one individual have nevertheless been disseminated across numerous blogs. There is obvious agreement among these blogs, and many people seem to interpret that consensus as proof that this individual's assertions are true. Yet this is a clear instance of *false* consensus.

One may hope that readers of these blogs take note of false consensus when they see it and properly discount the validity of information that is merely repeated. Alternatively, individuals may take repeated information just as seriously as information coming from different primary sources (e.g., the independent conclusions of many scientists that climate change is real). Might our sociological intuitions—like our scientific ones (Shtulman, 2017)—sometimes misguide us?

Here, we tested whether individuals do in fact distinguish between true and false consensus. We found that whereas individuals are sensitive to consensus (i.e., they are more likely to believe information corroborated by multiple sources), they fail to distinguish between true and false consensus, believing both equally. We then assessed whether individuals are sensitive to this difference when secondary sources are explicitly nonexperts (and could not have made any attempt to independently verify information). We also investigated whether there are any conditions under which people discount false consensus—and whether false-consensus errors can be explained by an explicit belief that a false consensus is just as valuable as a true consensus. Finally, we confirmed that these results are not due to a simple source-counting heuristic by showing that people correctly discount false consensus for information that can be directly perceived (e.g., witnessing an event). Data, materials, and preregistration information for all the experiments reported here can be found on the Open Science Framework (OSF) at https://osf.io/g9yrc.

Experiment 1: Assessing False Consensus

We first gauged sensitivity to the quality of consensus by showing participants fake news articles and measuring their confidence in the arguments presented. To manipulate consensus, we varied the number of secondary sources that took a particular side in a debate and the number of primary sources those secondary sources cited. (Note that *consensus*, as used here, means general agreement among some people rather than unanimous agreement.)

Method

Participants. Two hundred forty adult participants completed a survey online through Amazon Mechanical Turk (data from 22 additional participants who failed an attention check were discarded: 10 in the true-consensus condition, 4 in the false-consensus condition, 8 in the no-consensus condition; see the Procedure section). The sample size was chosen on the basis of independent pilot data and was preregistered. All participants lived in the United States.

Stimuli. Materials consisted of five fake news articles about the Japanese economy. Four of the articles took a positive stance (Japan's economy will continue to improve), and one took a negative stance (Japan's economy will not continue to improve). Participants were explicitly told that any markers of the articles' origins (as well as ads, etc.) had been removed to minimize distractions. All articles cited a single primary source (the name of an expert). To further emphasize the source, the expert's name was formatted to resemble a hyperlink (i.e., was written in blue underlined text; see https://osf.io/g9yrc for screenshots of the articles).

Procedure. Eighty participants were randomly assigned to three conditions: a true-consensus condition, a no-consensus condition, and a false-consensus condition. In the true-consensus condition, participants read all five articles (four positive, one negative). Because each article cited a unique primary source, participants heard from five primary sources. Participants in the false-consensus condition also read all five articles, except that the four positive articles all cited the same primary source. Thus, participants heard from only two primary sources overall (one taking a positive stance and one taking a negative stance). In the baseline no-consensus condition, participants read only two articles: the one negative article and one positive article randomly drawn from the four available positive articles. In this condition, then, participants heard from two primary sources (as in the false-consensus condition, one took a positive stance and one took a negative stance).

We expected that participants' confidence ratings would be significantly higher in the true-consensus condition than in the no-consensus condition. Therefore, the critical question was where confidence ratings in the false-consensus condition would fall. If people tracked only the number of primary sources, then confidence ratings in the false-consensus condition should be identical to those in the no-consensus condition, because in both, only a single primary source was cited on each side of the debate. On the other hand, if people's estimations of consensus relied only on the number of secondary sources, then confidence ratings in the false-consensus condition should be identical to those in the true-consensus condition, because both included four positive articles and one negative article.

Across all conditions, the exact same information was presented in each article; all that varied (other than superficial changes in the language) was the number of unique primary sources cited across the articles. The articles were presented in a randomized order, one at a time, to each participant. Participants had an unlimited amount of time to read each article before continuing.

After reading the articles, participants read the following prompt, presented in plain text at the top of the screen: "Based on the arguments you saw, to what extent do you agree that the economy in question will continue to improve? Please rate your agreement on the scale below from 0 (strongly disagree) to 100 (strongly agree)." Participants responded by clicking on a number line to indicate their confidence that the economy would improve and were prompted to confirm their answer before submitting it. On a separate screen, participants were then asked two questions. First, they saw a list of 10 sources and were asked to indicate which had been cited in the articles they read (there were no limits on the number of sources participants could select). Next, they saw a list of five nations and were asked to identify which one the articles had been about. Participants who answered incorrectly were excluded and replaced. No other information was collected.

Results

First, we assessed whether participants were tracking the individual sources. We used a d' analysis to determine whether participants correctly identified the sources they had seen and avoided sources they had not. In each condition, participants identified the correct primary sources at above-chance levels—true consensus: $d' = 1.16$, $t(79) = 8.63$, $p < .001$, $d = 0.97$; no consensus: $d' = 1.37$, $t(79) = 12.27$, $p < .001$, $d = 1.37$; false consensus: $d' = 2.20$, $t(79) = 14.21$, $p < .001$, $d = 1.60$. (For hit or false alarm rates of 0% or 100%, we used values of 5% and 95% instead. Modifying these values to be as conservative as possible did not change the results.) Thus, subsequent effects cannot be explained by failures to attend to (or remember) the sources in the first place.

The primary results of this experiment are shown in Figure 1. We first assessed whether consensus increased participants' confidence. Indeed, participants in the true-consensus condition were 15 points more confident (on a 100-point scale) than participants in the no-consensus condition, $t(158) = 4.71$, $p < .001$, $d = 0.74$, Bonferroni corrected. But did participants discount the value of consensus in the false-consensus condition? First, we assessed whether confidence ratings in the true-consensus condition differed from those in the false-consensus condition. The two were separated by only 2 points on a 100-point scale, and this difference was not significant, $t(158) = 0.69$, $p > .50$, $d = 0.12$, Bonferroni corrected. Additionally, we tested whether confidence in the false-consensus condition was greater than in the no-consensus condition. Indeed, participants were more confident in the false-consensus condition than in the no-consensus condition, amounting to a 13-point difference in confidence, or about a 23% increase, $t(158) = 4.02$, $p < .001$, $d = 0.61$, Bonferroni corrected.

In principle, this effect could have been driven by participants who failed to notice there was a repeated source in the false-consensus condition. But in fact, participants across all conditions were quite good at identifying which sources they had encountered. Furthermore, average confidence ratings in the true- and false-consensus conditions did not differ from one another. If the similarity of the true- and false-consensus conditions was driven by only a few participants, one might expect this difference to be more pronounced. To test this empirically, we removed participants whose d' scores were not higher than zero (i.e., participants who were particularly poor at identifying the sources they had heard from). Doing so actually made the effect even greater (likely because of the removal of extreme scores and noise): The difference between false consensus and no consensus increased to 16 points (amounting to more than a 30% increase in confidence).[1] In practice, we tested many different criteria for inclusion. Even when we took the most extreme criteria—including only participants who identified every single source correctly and no sources incorrectly—the pattern remained the same: Confidence ratings for both the

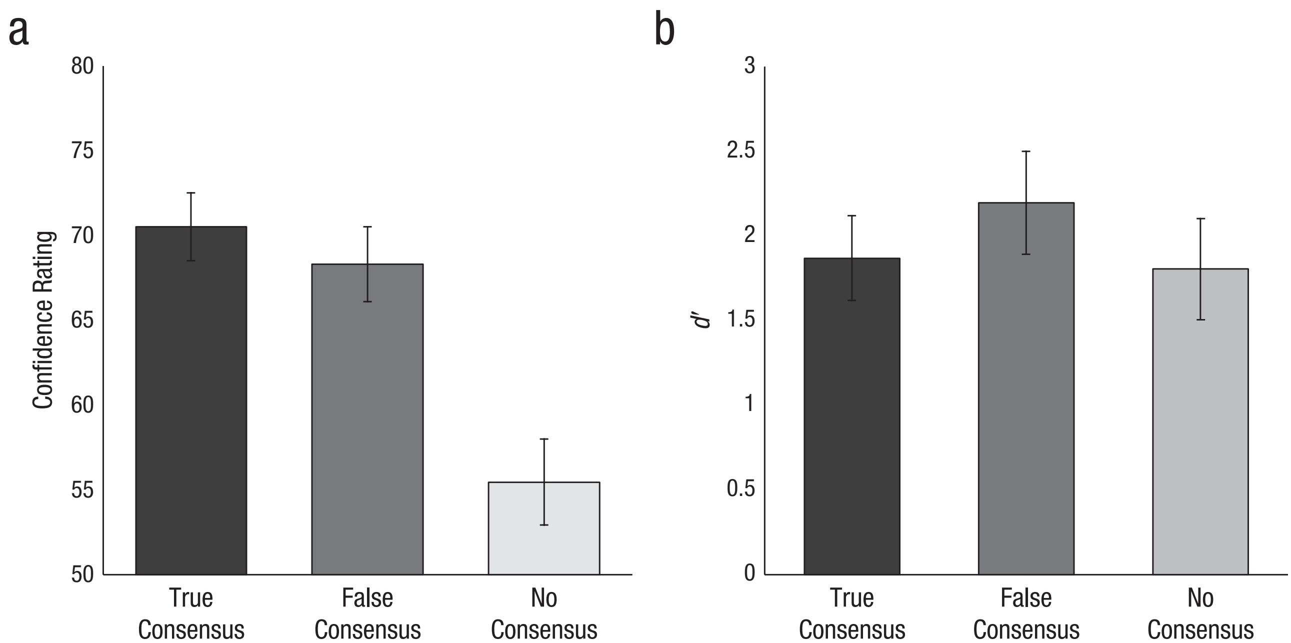

Fig. 1. Results from Experiment 1: (a) mean confidence rating and (b) mean d' score in the true-, false-, and no-consensus conditions. Error bars represent ±1 *SE*.

true-consensus and false-consensus conditions were greater than those in the no-consensus condition (ps < .05) but no different from each other (p = .62).[2]

Discussion

The results of Experiment 1 suggest that individuals are not sensitive to any difference between true and false consensus; however, there may be a rational explanation for the results. When reading news articles, it is possible that individuals assume that news sources possess some amount of journalistic integrity, evaluating their sources before citing them. Thus, if many journalists all talk to one source, individuals might assume that this source is highly qualified.

Experiment 2: Minimizing Expertise

We next assessed sensitivity to consensus in a case in which participants were unlikely to make assumptions about the knowledge and expertise of the secondary sources. Participants read fake student essays rather than news articles. The purpose of this manipulation was to make clear to participants that the secondary sources had not—and indeed could not have—independently verified the claims they were asserting (because they lacked expertise). In addition, we ensured that the information participants received across stimuli was matched such that participants were exposed to exactly the same information in both the false-consensus and no-consensus conditions.

Method

All elements of the experimental design were identical to those of Experiment 1, except as stated below. Two hundred forty new participants completed the survey online through Amazon Mechanical Turk (data from 59 additional participants who failed a simple attention check were discarded: 22 in the true-consensus condition, 16 in the false-consensus condition, 21 in the no-consensus condition). This sample size was chosen to be identical to that in the previous experiment. This experiment was also preregistered.

Instead of reading news articles about the state of the Japanese economy, participants read fake student essays about a proposed tax policy in Sweden. Prior to reading the articles, participants were told explicitly (in bold letters and at the center of the screen) that the students had been specifically instructed to cite their sources in order to make their arguments. Unlike the sources in the previous experiment, the primary sources in this experiment were the names of actual economic foundations. As before, four articles took a positive stance (the tax policy should be approved), and one article took a negative stance (the tax policy should not be approved). Participants read either all of the articles (true- and false-consensus conditions) or only one article on each side (no-consensus condition). To equate the information participants received across conditions, we designed the positive articles to contain six relevant facts. Each of the articles made reference to these facts (albeit in a different order and using slightly different language) and provided no other information. This ensured that any difference between the false-consensus and no-consensus conditions was not due to a difference in information.

Results

Again, we asked whether participants were tracking the individual sources cited in the essays. The results of our d' analysis suggest that participants were closely tracking sources. In each condition, participants identified the correct primary sources at above-chance levels—true consensus: $d' = 1.96$, $t(79) = 16.20$, $p < .001$, $d = 1.81$; no consensus: $d' = 1.92$, $t(72) = 10.41$, $p < .001$, $d = 1.16$; false consensus: $d' = 2.05$, $t(79) = 12.36$, $p < .001$, $d = 1.38$.

The primary results of this experiment are shown in Figure 2. Analyses confirmed what is evident from the figure. Consensus increased participants' confidence in the information they read, with participants in the true-consensus condition 12 points more confident (on a 100-point scale) than participants in the no-consensus condition, $t(158) = 3.67$, $p < .001$, $d = 0.55$, Bonferroni corrected. However, the mean ratings in the true-consensus and false-consensus conditions were separated by only 1 point on a 100-point scale, and this difference was not significant, $t(158) = 0.33$, $p > .90$, $d = 0.06$, Bonferroni corrected. Participants were also more confident in the false-consensus condition than in the no-consensus condition, amounting to an 11-point difference in confidence, $t(158) = 3.34$, $p = .003$, $d = 0.50$, Bonferroni corrected.

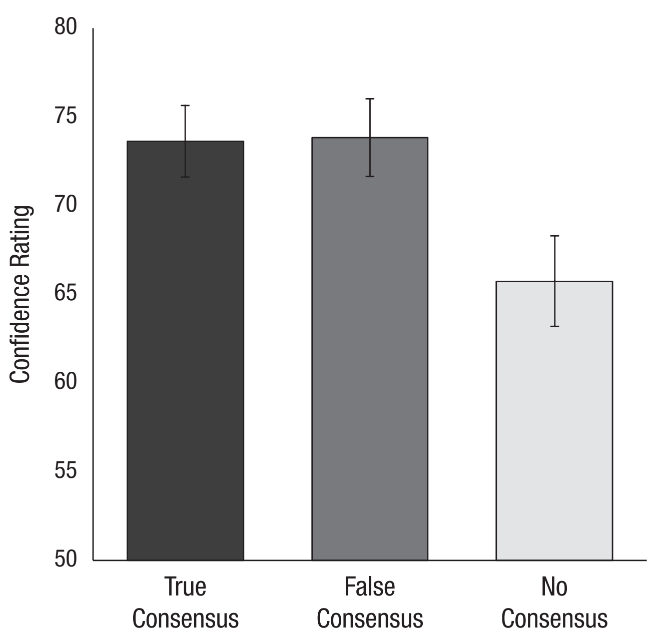

Fig. 2. Results from Experiment 2: mean confidence rating in the true-, false-, and no-consensus conditions. Error bars represent ±1 *SE*.

We again tested whether these results could be driven by a subset of the participants who did not attend to source information. After removing participants with d' scores less than or equal to zero, the difference between the true-consensus and no-consensus conditions was still 11 points, and the difference between the false-consensus and no-consensus conditions was still 9 points.

Discussion

Experiment 2 provided converging evidence that individuals do not properly discount information that is merely repeated. Notably, confidence ratings were higher overall than those in Experiment 1. This is likely due to a difference in the stimuli; in this experiment, we may have inadvertently made the positive side of the debate more persuasive. Importantly, however, the key contrast here is the relative difference between the true-consensus and no-consensus conditions.

Experiment 3: Highlighting Sources

How robust is this illusion of consensus? In the first two experiments, we showed that participants' failure to distinguish between true and false consensus could not be explained by a failure to notice or remember different sources in the first place. However, participants were probed about sources after making their confidence judgments. Would individuals differentiate true and false consensus if explicitly probed about sources before making a confidence judgment?

Method

All elements of the experimental design were identical to those of Experiment 1, except that participants made their confidence judgments after being asked to identify the sources they had heard from. Two hundred forty new participants were recruited through Amazon Mechanical Turk (data from 15 additional participants who failed an attention check were discarded: 5 in the true-consensus condition, 2 in the false-consensus condition, 8 in the no-consensus condition). This sample size was chosen to be identical to those in the previous experiments. This experiment was also preregistered.

Results

Again, we found that participants were attending to the primary sources. In each condition, participants identified the correct primary sources at above-chance levels—true consensus: $d' = 1.90$, $t(79) = 14.76$, $p < .001$, $d = 1.65$; no consensus: $d' = 1.31$, $t(79) = 10.78$, $p < .001$, $d = 1.21$; false consensus: $d' = 2.24$, $t(79) = 16.64$, $p < .001$, $d = 1.86$.

Consensus did increase participants' confidence in the information they read: Participants in the true-consensus condition were 14 points more confident (on a 100-point scale) than participants in the no-consensus condition, $t(158) = 4.19$, $p < .001$, $d = 0.61$, Bonferroni corrected. Again, we also assessed whether confidence ratings in the true-consensus condition differed from those in the false-consensus condition. The two were separated by 3 points on a 100-point scale, and this difference was not significant, $t(158) = 0.99$, $p > .50$, $d = 0.17$, Bonferroni corrected. Participants were also more confident in the false-consensus condition than in the no-consensus condition, amounting to an 11-point difference, or about a 19% increase in confidence, $t(158) = 3.20$, $p = .004$, $d = 0.51$, Bonferroni corrected.

As in the previous experiments, we tested whether these effects could be driven by some subset of the participants who did not attend to source information. After removing participants with d' scores less than or equal to zero, the effects were once again even stronger: The difference between the true-consensus and no-consensus conditions increased to 15 points, and the difference between the false-consensus and no-consensus conditions increased to 12 points.

However, in a separate analysis, we excluded the upper and lower quintiles of the distributions of confidence judgments in each condition (noting that many people seemed to respond with unusually extreme scores). The resulting sample told a slightly different story: Although confidence ratings in the false-consensus condition were still significantly higher than those in the no-consensus condition, $t(94) = 5.70$, $p < .001$, $d = 1.04$, Bonferroni corrected, they were lower than those in the true-consensus condition, $t(94) = 2.84$, $p = .01$, $d = 0.79$, Bonferroni corrected. Thus, participants may have partially discounted false consensus. The results after excluding the upper and lower quintiles of the distributions can be seen in Figure 3. (We tried many possible exclusion criteria for this analysis, and there is nothing special about the removal of quintiles. We chose this value because it seemed like a principled middle ground: We removed enough of the data to ensure that all extreme scores were removed but retained enough that our samples were still rather large and the data were clearly interpretable.)

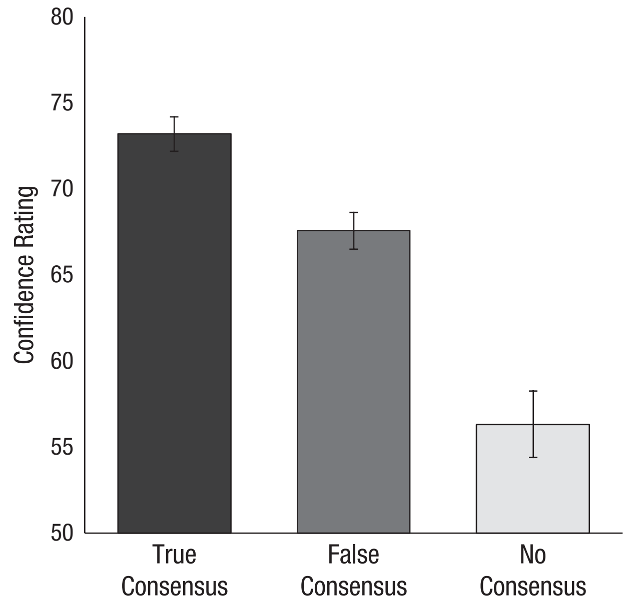

Fig. 3. Results from Experiment 3: mean confidence rating in the true-, false-, and no-consensus conditions after the upper and lower quintiles of each distribution were removed. Error bars represent ±1 *SE*.

We took several steps to ensure that the above analysis was valid and that its results were not spurious. First, we retroactively analyzed the data from Experiments 1 and 2 this way and found that in both cases, excluding participants in the same way produced results that were actually stronger (i.e., the differences between the no-consensus condition and the other two conditions grew, and the difference between the true- and false-consensus conditions shrank). Second, we ran a preregistered replication of this exact experiment including this additional analysis. The results were identical. The preregistration for this replication, as well as the full data from this experiment, can be found on the project's OSF page (https://osf.io/g9yrc).

Discussion

Participants in this experiment were explicitly instructed (in bold type) to attend to each article's sources and cued to reflect on these sources (by recalling them before making their confidence judgments). Despite these reminders, participants were still deceived by false consensus.

Experiment 4: Explicit Comparison

Across three experiments, we showed that failure to distinguish between true and false consensus cannot be explained by a failure to notice different sources. However, these results could stem either from an illusion of consensus or from an explicit belief that a false consensus is valuable. Experiment 4 probed such explicit beliefs.

Method

All elements of the experimental design were identical to those of Experiment 1, except as noted below. Two hundred forty new participants were recruited through Amazon Mechanical Turk (data from 39 participants who failed an attention check were discarded: 23 in the true-consensus condition, 10 in the false-consensus condition, 6 in the no-consensus condition). This sample size was chosen to be identical to those in the previous experiments. This experiment was also preregistered.

Before completing the task, participants were asked to consider a hypothetical scenario in which they read several news articles. They were asked to contrast a case in which those news articles all cited unique primary sources to a case in which those news articles cited the same primary source. They were told, "The articles you read seem to mostly agree. Here's the question: as you're reading, you notice that each article cites its own **unique** primary source. Or, you notice that each article cites **the same** primary source." They then had the option of indicating that the former was more believable, the latter was more believable, or that both were equally believable. Participants were reminded that there was no correct answer and that "we [were] just interested in what people think."

Results

Overall, 50% of the participants believed that a true consensus was most believable, 16% believed that a false consensus was most believable, and 33% believed that both were equally believable. Critically, the preference for true consensus was significant ($p < .001$). There were no differences across the conditions.

We first analyzed the results of only participants who rated true consensus as more believable. Participants correctly identified primary sources at above-chance levels (*p*s < .001). Second, although confidence judgments in the true- and false-consensus conditions were greater than in the no-consensus condition—20 points: $t(75) = 4.37$, $p < .001$, $d = 1.05$; 14 points: $t(84) = 0.29$, $p = .004$, $d = 0.67$, respectively, Bonferroni corrected—the two conditions did not differ significantly from one another, $t(77) = 1.28$, $p > .40$, $d = 0.30$, Bonferroni corrected. The results were identical when data from all 80 participants were included, when we excluded subzero d' scores, and when we excluded the upper and lower quintiles (per the analysis in Experiment 3). In short, participants still fell prey to an illusion of consensus, even after explicitly identifying this information as less believable.

Discussion

It appears that the failure to distinguish between true and false consensus cannot be explained by an explicit belief. Even though large numbers of participants indicated that true consensus was not superior to false consensus, the illusion of consensus was apparent even when we restricted our analyses to those participants who indicated the contrary: that true consensus was superior to false consensus. This finding makes the previous results all the more surprising: Individuals believe that a false consensus should be discounted yet fail to do so.

Experiment 5: Bear Sighting

What is the nature of the consensus illusion? Do people always attend solely to the number of primary sources? Or are they susceptible to an illusion of consensus only in domains that possess some degree of unknowability? In Experiment 5, we examined whether participants still fail to discount false consensus when they clearly should: when only one person directly perceived an event.

Method

All elements of the experimental design were identical to those of Experiment 1, except that the vignettes were fake news articles about a bear sighting on a local high school campus. The relevant sources were students at the school who claimed that a bear either was or was not present. Unlike in the other experiments, the claim in question (i.e., whether or not a bear was sighted) is very knowable in the sense that anyone could see a bear if one were present; no expertise would be required. This is different from the design of our other experiments, which required making inferences from data. Two hundred forty new participants were recruited through Amazon Mechanical Turk (data from 27 additional participants who failed an attention check were discarded: 15 in the true-consensus condition, 7 in the false-consensus condition, 5 in the no-consensus condition). This sample size was chosen to be identical to those in the previous experiments. This experiment was also preregistered.

Results

Again, we found that participants were attending to the primary sources. In each condition, participants identified the correct primary sources at above-chance levels—true consensus: $d' = 1.18$, $t(79) = 10.47$, $p < .001$, $d = 1.08$; no consensus: $d' = 1.50$, $t(79) = 9.82$, $p < .001$, $d = 1.07$; false consensus: $d' = 1.72$, $t(79) = 12.28$, $p < .001$, $d = 1.32$.

Consensus did increase participants' confidence in the information they read: Participants in the true-consensus condition were 23 points more confident (on a 100-point scale) than participants in the no-consensus condition, $t(158) = 5.89$, $p < .001$, $d = 0.95$, Bonferroni corrected. Again, we also assessed whether confidence ratings in the true-consensus condition differed from those in the false-consensus condition. Note that here we did expect participants to appropriately discount false consensus because only one person saw the bear in the false-consensus condition. Indeed, that is exactly what we found: The two were separated by 19 points on a 100-point scale, $t(158) = 4.82$, $p < .001$, $d = 0.77$, Bonferroni corrected. Only 4 points separated participants in the false-consensus and no-consensus conditions, and there was no difference between the two, $t(158) = 1.07$, $p > .50$, $d = 0.17$, Bonferroni corrected.

However, following our procedure in Experiment 3 (and consistent with our preregistration for this experiment), we excluded the upper and lower quintiles of the distributions of confidence judgments in each condition. The resulting sample told a slightly different story: Although confidence ratings in the false-consensus condition were still significantly lower than those in the true-consensus condition, $t(94) = 7.77$, $p < .001$, $d = 1.72$, Bonferroni corrected, they were slightly higher than those in the no-consensus condition, $t(94) = 2.77$, $p = .019$, $d = 0.50$, Bonferroni corrected. Thus, participants still may not have fully discounted false consensus.

Discussion

Here, we tested whether people discount false consensus when testimony is based on a directly perceived event. Whereas people might not always fully discount false consensus, they do substantially discount it, which suggests that the illusion documented in the prior experiments does not stem merely from a source-counting heuristic. The phenomenon seems most robust for relatively unknowable information.

General Discussion

In Experiment 1, individuals were just as likely to believe information derived from a true consensus as from a false consensus. In Experiment 2, this effect persisted even when the expertise of the secondary sources was minimized; in Experiment 3, this effect persisted even when individuals were explicitly cued to attend to primary sources. Experiment 4 demonstrated that these findings cannot be explained by explicit beliefs that true and false consensus are equally valuable. Finally, Experiment 5 showed that this illusion of consensus is not merely the product of a simple source-counting heuristic: This phenomenon may be most robust in cases in which information is relatively unknowable. In sum, we found that individuals are often sensitive to consensus in a superficial rather than informative manner.

Misinformation in context

How far reaching is this illusion of consensus? One salient real-world example comes from science-denying climate-change blogs, which overwhelming rely on a single, unaccredited source (Harvey et al., 2018). Yet this phenomenon likely applies in a much broader range of situations: from interpreting rumors in the workplace, to identifying trustworthy news sources, to reasoning about empirical research, to inferring the popularity of opinions, and beyond (e.g., Weaver, Garcia, Schwarz, & Miller, 2007). More generally, we think these findings contribute to a broader science of misinformation (e.g., Hahn, Harris, & Corner, 2016; Kahan, 2017; Pennycook & Rand, 2018; Shtulman, 2017), helping us understand how misinformation becomes widespread.

However, some unanswered questions remain. For example, it is not entirely clear why participants fail to discount false consensus. Experiment 2 was designed to minimize assumptions made about the secondary sources (students), but participants could still reasonably assume that students acted as some kind of filter. For example, perhaps students all attended to (and wrote about) the same primary source because that information was most persuasive. This is consistent with work suggesting that people make rich inferences not only about the content of sources but also about other source characteristics (e.g., whether they have a prior history of being correct or whether they are an authority on the subject; see Collins, Hahn, von Gerber, & Olsson, 2018). In the present work, although a clear majority of participants in Experiment 4 preferred true consensus, a sizable proportion believed that a false consensus was equally or more valuable. But what specific inference causes this preference for false consensus? Future work may address this question by probing participants' explicit beliefs or by studying the phenomenon of false consensus in instances in which the secondary sources could not possibly have outside knowledge.

These effects might also vary in robustness across contexts. Our examples purposefully straddled the fence between objectivity and subjectivity. Arguments of taxation and economics are grounded in empirical data, yet there is room for interpretation. Other work has shown that people interpret consensus differently depending on the level of perceived subjectivity or objectivity of information (Yousif & Keil, 2018). Such factors might also play a role as people interpret the quality of a consensus.

Conclusion

As we try to decide what information to trust, false consensus is a powerful cue. The results of the present experiments bear directly on everyday behavior: For example, our work helps explain how misinformation becomes widely propagated despite limited evidence (e.g., Harvey et al., 2018; Pennycook & Rand, 2018; Shtulman, 2017). How can we combat this overreliance on consensus? Should news sources explicitly differentiate between claims that are being repeated and claims that have been verified, and would this even make a difference? Results of Experiment 3 suggest that intervention is not hopeless and that individuals can be made to discount false consensus when very explicitly cued to source information. Understanding this illusion of consensus—and how it can be combatted—is critical in an era of social media, rapid news cycles, and increasingly polarized ideological echo chambers.

Action Editor

Marc J. Buehner served as action editor for this article.

Author Contributions

S. R. Yousif and R. Aboody conceived and designed the study. Testing, data collection, and data analysis were performed by S. R. Yousif with input from R. Aboody and F. C. Keil. S. R. Yousif drafted the manuscript, and R. Aboody and F. C. Keil provided critical revisions. All authors approved the final version of the manuscript for submission.

Acknowledgments

For helpful comments and conversation, we thank Woo-kyoung Ahn, Brynn Sherman, Clara Colombatto, Richard Ahl, Babak Hemmation, and all the members of the Yale Cognition and Development Lab.

Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

Open Practices

TC

All data and materials have been made publicly available via the Open Science Framework and can be accessed at https://osf.io/g9yrc. The design and analysis plans for the experiments were preregistered at aspredicted.org, and copies of the preregistrations can be found at https://osf.io/g9yrc (additional nonpreregistered analyses were conducted for some experiments, as noted in the text, but they did not change the results overall). The complete Open Practices Disclosure for this article can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797619856844. This article has received the badges for Open Data, Open Materials, and Preregistration. More information about the Open Practices badges can be found at http://www.psychologicalscience.org/publications/badges.

Notes

1. This specific analysis was not preregistered, nor was the specific exclusion criterion. However, we did preregister our intention to collect accuracy scores and to demonstrate that participants were paying attention. This is likewise true for all subsequent experiments.
2. Because our results remained unchanged either way, we conducted this analysis using only above-zero d' scores in subsequent experiments.

References

Asch, S. E. (1956). Studies of independence and conformity: I. A minority of one against a unanimous majority. *Psychological Monographs: General and Applied*, *70*, 1–70.

Burdett, E. R. R., Lucas, A. J., Buchsbaum, D., McGuigan, N., Wood, L. A., & Whiten, A. (2016). Do children copy an expert or a majority? Examining selective learning in instrumental and normative contexts. *PLOS ONE*, *11*(10), Article e0164698. doi:10.1371/journal.pone.0164698

Collins, P. J., Hahn, U., von Gerber, Y., & Olsson, E. J. (2018). The bi-directional relationship between source characteristics and message content. *Frontiers in Psychology*, *9*, Article 18. doi:10.3389/fpsyg.2018.00018

Corriveau, K. H., Fusaro, M., & Harris, P. L. (2009). Going with the flow: Preschoolers prefer nondissenters as informants. *Psychological Science*, *20*, 372–377.

Corriveau, K. H., & Harris, P. L. (2010). Preschoolers (sometimes) defer to the majority in making simple perceptual judgments. *Developmental Psychology*, *46*, 437–445.

Einav, S. (2014). Does the majority always know best? Young children's flexible trust in majority opinion. *PLOS ONE*, *9*(8), Article e104585. doi:10.1371/journal.pone.0104585

Einav, S. (2018). Thinking for themselves? The effect of informant independence on children's endorsement of testimony from a consensus. *Social Development*, *27*, 73–86.

Fusaro, M., & Harris, P. L. (2008). Children assess informant reliability using bystanders' non-verbal cues. *Developmental Science*, *11*, 771–777.

Hahn, U., Harris, A. J., & Corner, A. (2016). Public reception of climate science: Coherence, reliability, and independence. *Topics in Cognitive Science*, *8*, 180–195.

Harvey, J. A., van den Berg, D., Ellers, J., Kampen, R., Crowther, T. W., Roessingh, P., . . . Mann, M. E. (2018). Internet blogs, polar bears, and climate-change denial by proxy. *BioScience*, *68*, 281–287. doi:10.1093/biosci/bix133

Hu, J. C., Buchsbaum, D., Griffiths, T. L., & Xu, F. (2013). When does the majority rule? Preschoolers' trust in majority informants varies by domain. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Meeting of the Cognitive Science Society* (pp. 2584–2589). Austin, TX: Cognitive Science Society.

Hu, J. C., Whalen, A., Buchsbaum, D., Griffiths, T. L., & Xu, F. (2015). Can children balance the size of a majority with the quality of their information? In D. C. Noelle, R. Dale, A. S. Warlaumont, J. Yoshimi, T. Matlock, C. D. Jennings, & P. P. Maglio (Eds.), *Proceedings of the 37th Annual Meeting of the Cognitive Science Society* (pp. 956–961), Austin, TX: Cognitive Science Society.

Kahan, D. M. (2017). ‘Ordinary science intelligence’: A science-comprehension measure for study of risk and science communication, with notes on evolution and climate change. *Journal of Risk Research*, *20*, 995–1016.

Kundey, S. M., German, R., De Los Reyes, A., Monnier, B., Swift, P., Delise, J., & Tomlin, M. (2012). Domestic dogs’ (*Canis familiaris*) choices in reference to agreement among human informants on location of food. *Animal Cognition*, *15*, 991–997.

Pennycook, G., & Rand, D. G. (2018). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*, *188*, 39–50.

Ross, L., Greene, D., & House, P. (1977). The “false consensus effect”: An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology*, *13*, 279–301.

Shtulman, A. (2017). *Scienceblind: Why our intuitive theories about the world are so often wrong*. New York, NY: Basic Books.

Weaver, K., Garcia, S. M., Schwarz, N., & Miller, D. T. (2007). Inferring the popularity of an opinion from its familiarity: A repetitive voice can sound like a chorus. *Journal of Personality and Social Psychology*, *92*, 821–833.

Yousif, S. R., & Keil, F. C. (2018). When in doubt: Using confidence and consensus as ‘summary statistics’ of collective knowledge. In T. T. Rogers, M. Rau, X. Zhu, & C. W. Kalish (Eds.), *Proceedings of the 40th Annual Conference of the Cognitive Science Society* (pp. 2741–2746). Austin, TX: Cognitive Science Society.