

Inconsistent Arguments are Perceived as Better Than Appeals to Authority: An Extension of the Everyday Belief Bias

Calvin Deans-Browne (cdeansbrowne@gmail.com)
Alexandra Băitanu (alexandra.baitanu.20@ucl.ac.uk)
Yuliya Dubinska (jldubinsky08@gmail.com)
Henrik Singmann (hsingmann@gmail.com)

Department of Experimental Psychology, University College London, 26 Bedford Way,
London, WC1H 0AP England

Abstract

Social media is often used as a platform where individuals engage in debate regarding topics that are important to them. Not all arguments are equally convincing, and whilst a given argument may be persuasive to some people, it is often seen as inadequate by others. We are interested in both the individual and argument level differences that make ‘everyday’ arguments such as those on social media persuasive. In a replication of our Everyday Belief Bias Task (Deans-Browne & Singmann, 2024), we investigate this question using a paradigm that consists of two parts. In the first part, we measure participant’s individual beliefs about eight claims each referring to a political topic (e.g., *Abortion should be legal*). In the second part, participants rated an argument for each of these claims that was deemed as either *good*, *inconsistent* (containing internal inconsistencies), or *authority-based* (being centered around appeals to authority). We replicated the belief consistency effect – participants preferred arguments that were also in line with their beliefs. We also found that authority-based arguments were rated as worse than inconsistent arguments, and that both types of arguments were rated as worse than good arguments. The implications are first that people do not evaluate arguments independently of the background beliefs held about them. Secondly, people are willing to ignore inconsistencies in arguments more than they are willing to accept the endorsement of authority figures as adequate evidence for arguments.

Keywords: belief bias; argument quality; social media; informal reasoning

Introduction

In the digital age, we are often presented with information from sources of varying credibility. This has made media literacy a skill of increasing importance, even more so as the algorithms designed to increase user engagement on social media can also spread misinformation (e.g., Menczer, 2021; Gogarty et al., 2023). Research into media literacy largely falls into two areas; the type of information people are exposed to (e.g., Bakshy et al., 2015; Cinelli et al., 2021) and people’s ability to recognize misinformation when exposed to it (e.g., Pennycook and Rand, 2019). Recently, we connected a finding from the traditional reasoning literature, the belief bias (e.g., Evans et al., 1983), with research on media literacy. Our findings showed that people’s perception of everyday political arguments, as found on social media, was biased by their prior beliefs (Deans-Browne & Singmann, 2024).

In the *Everyday Belief Bias Task*, we showed participants a series of arguments that varied in terms of their informal argument quality. The arguments participants saw were either *good* or *bad*. *Good* arguments contained evidence that strongly supported their conclusions. In contrast, *bad* arguments contained common argument fallacies (circularity, appeals to authority, appeals to popularity) and only weak evidence in support of their conclusions. We found that whilst participants generally rated good arguments as better than bad ones, they also tended to rate arguments in line with their beliefs as better than arguments that were not. These results are analogous to those found in the traditional belief bias literature (e.g., Evans et al., 1983; Klauer et al., 2000; Trippas et al., 2018) in which arguments vary by their formal argument quality (i.e., valid vs invalid). Importantly, whereas in the traditional belief bias paradigm beliefs are generally indisputable (i.e., everyone agrees on the same beliefs), we investigated *disputable* beliefs revolving around contentious political topics (e.g., abortion, climate change).

One feature of the *Everyday Belief Bias Task* is that the operationalization of good and bad arguments is inherently ambiguous. As the focus is on informal argumentation, the criterion for which an argument should be labelled as good or bad is unclear. The goal of the present work is to address the source of ambiguity and investigate which aspect of an argument makes it perceived as bad.

In our previous experiment (Deans-Browne & Singmann, 2024), the good arguments for each topic were based on already established arguments prevalent in the current discourse. The arguments were inspired by those from www.procon.org, a non-partisan website that aims to keep users informed on various topics by offering both sides of the political debate. Thus, what made a good argument informally “good” was fixed by the topic of the argument itself. The good arguments therefore presented the best available evidence for either side of the claim.

In contrast, bad arguments were not based on evidence, but on various reasoning fallacies (circularity, appeal to authority, appeal to popularity). However, the choice of fallacy for each topic was not done systematically. In the present paper we are interested in comparing the perceived ‘badness’ of two different reasoning fallacies, inconsistent arguments and arguments based on appeal to authority, in a systematic manner.

Argument Quality of Informal Arguments

For formal arguments, such as the syllogisms used in traditional belief bias studies, argument quality is independent of the content of the argument and solely determined by an argument’s form. In contrast, for the informal arguments investigated here, the quality is solely determined by its content. All informal arguments are formally invalid (i.e., the conclusion does not follow from the premises with logical necessity).

There have been some studies investigating informal arguments of this type. For example, Stanovich and West (1997) developed the Argument Evaluation Test which asked participants to judge the quality (strength) of arguments presented as part of a conversation. Each argument consisted of a single sentence that was part of the conversational dialogue. The informal argument strength was not systematically manipulated but instead based on the author’s intuition and later confirmed through the ratings of eight philosophy experts. In line with the belief bias literature, they found that participants’ beliefs predicted perceived argument quality more strongly than objective argument quality (see Thompson & Evans, 2012, for a replication of this pattern).

Another study investigating informal arguments (also in the context of the belief bias) is Wolfe and Kurby (2017). Like Stanovich and West (1997), their arguments consist of a single sentence containing both a premise and conclusion. In their study, good and bad arguments were more systematically varied. For good arguments the premise was relevant to the conclusion whereas for bad arguments it was not. Participants had no problem distinguishing between good and bad arguments, but their quality ratings were again also affected by their prior beliefs.

Informally good and bad arguments more like those found in the real world (i.e., consisting of more than one sentence) were investigated by Hopkins et al. (2016). Their interest was whether reductive arguments (i.e., those referring to more fundamental processes) are perceived as better than arguments of similar quality without reductive information. Importantly, in addition to manipulating the presence of reductive information they also introduced a systematic manipulation of informal argument quality. For the construction of good arguments, “researchers or textbooks

provided for the phenomena, and our expert consultants verified that these [arguments] were clear and accurate” (p. 70). In contrast, bad arguments “were constructed by removing the key explanatory information from the good [arguments] and replacing it with either circular restatements of the phenomenon or with non-explanatory information” (p. 70). Participants in their study were clearly able to distinguish between the good and bad arguments.

Finally, a study by Harris and colleagues (2016) explicitly investigated people’s perception of different types of arguments based on appeal to authority. Their results found that the perceived quality of an argument from authority depends on the trustworthiness and expertise of the authority figure in line with a Bayesian account of argumentation. Unfortunately, their study did not compare arguments from authority with other types of arguments.

This brief survey of the literature demonstrates that until recently, informal arguments studied in the psychological literature have typically been impoverished in nature, consisting only of one or two sentences (i.e., premise plus conclusion). The literature also suggests that people have no problems distinguishing good from bad arguments, especially if the bad arguments consist of a premise that is irrelevant or incompatible with its conclusion. In recent years, richer arguments have been studied, but a comparison of different types of bad arguments appears to be missing from the literature as of yet.

The Everyday Belief Bias Task

The *Everyday Belief Bias Task* (Deans-Browne & Singmann, 2024) consists of two parts. The first part of the study consists of measuring participants’ beliefs for the eight political claims shown in Table 1. Participants are asked to rate the veracity of each claim on a 7-point Likert scale with anchors ranging from *extremely false* to *extremely true*. The claims were associated with *disputable beliefs*; for each claim, some participants strongly believed it to be true, whereas other participants strongly believed it to be false.

In the second part, participants were asked to rate the quality of arguments relating to these claims on a 6-point Likert scale with anchors ranging from *extremely bad* to *extremely good*. For each claim, participants saw exactly one

Table 1: List of topics and associated claims

| Topic | Claim |
|-------------------------------------|--|
| Climate change | Human activity is primarily responsible for climate change |
| Abortion | Abortions should be legal in the US. |
| Kneeling during the national anthem | Kneeling during the national anthem is an appropriate form of protest. |
| Private prisons | Private prisons are not well run. |
| Fracking | It is in the United States’ best interest to continue fracking. |
| Habitual offenders | Habitual offender (or “three strike”) laws are an appropriate way to punish reoffenders. |
| Gun control | Further gun control laws are unnecessary. |
| Cancel culture | Cancel culture is bad for society. |

Table 2: Examples of good, inconsistent, and authority-based arguments

| Argument type examples |
|--|
| Pro-good argument Abortions under <i>Roe v. Wade</i> balanced two fundamental rights; the right of the pregnant woman to bodily autonomy and the right of the unborn child to life. The unborn child only has the potential for life as we know it when they can survive outside the womb, and abortions had to occur before this stage under this ruling. Consequently, abortions can be consistent with both fundamental rights. Abortion should therefore be legal in the US. |
| Pro-inconsistent argument Abortions are safe procedures that protect lives. Women that are denied abortions are likely to later have poorer mental and physical health, alongside financial problems. Abortion promotes the idea that human lives are disposable when inconvenient. This has worryingly advanced into decisions to terminate lives of fetuses with unwanted characteristics like disabilities or a specific sex. Abortion protects the bodily autonomy of women – a fundamental human right. Therefore, abortions should remain legal in the US. |
| Pro-authority argument Abortions should be legal because they have been performed for a long time. Only recently, the protective <i>Roe v. Wade</i> law was overturned. Now, each state can impose their own anti-abortion laws. Current President Joe Biden voiced his beliefs that women’s rights for abortion must be protected. He has also argued against the recent overturn. He is the current President of the United States and pro abortions. Therefore, abortions should remain legal in the US. |

argument. The eight arguments participants saw were manipulated across two factors, argument support and argument quality. Argument support referred to whether the argument was in line with or in opposition to the claim; half the arguments participants saw were in line with the claims they had rated (referred to as *pro* arguments), while the other half were in opposition to the remaining claims they had rated (referred to as *con* arguments). Argument quality referred to whether the informal quality of an argument was good or bad as described above.

As mentioned previously, our results were largely analogous to the pattern observed in the traditional belief bias paradigm. In addition to the main effect of argument quality – participants were able to distinguish between good and bad arguments – we also found an argument support by belief interaction. This interaction revealed a belief consistency effect for participants’ argument quality ratings. For *pro* arguments there was a positive relationship between participants’ beliefs and their argument quality ratings, whereas for *con* arguments this relationship was negative. Importantly, the overall magnitude of the belief consistency effect was around 2.5 times larger than the effect of argument

quality. In contrast to the traditional belief bias paradigm where a belief by (formal) argument quality is regularly observed, we found no evidence for the belief by (informal) argument quality interaction (in line with Wolfe & Kurby, 2017).

The Current Study

The goal of the current study was to compare informal arguments that were either good, inconsistent, or authority-based. Examples can be seen in Table 2. Good arguments, as described previously, were based on arguments that are already in the political discourse for the topic in question (see Deans-Browne & Singmann, 2024).

Inconsistent arguments were based on the same contents as the *pro* and *con* good arguments, but their configuration was not internally consistent, such that each argument always contained parts from both *pro*-good and *con*-good arguments. Statements at the beginning and end of the arguments were based on arguments that supported the overall conclusion, and a statement in the middle of the inconsistent argument was based on arguments that went against the overall conclusion. In this way, the inconsistent arguments were made up of statements supporting the conclusion which ‘sandwiched’ a statement against it.

Authority-based arguments, on the other hand, emphasized the endorsement of an authority figure as evidence of its conclusion. The authority figures have expressed the views written in the argument in some form of media. In none of the arguments did we include any incorrect information to the best of our knowledge.

Methods

Participants

Participants were recruited through Prolific and restricted to native English speakers in the USA. A total of 187 participants took part in the study. Of those, 34 failed attention checks and one did not speak English natively (a pre-requisite for participation). This left us with 152 participants (72 male, 78 female, 2 other) from whom we analyzed data. With regards to their highest education levels; one had not completed high school, 21 had completed high school, 34 had completed some college, eight were currently in college, 57 had completed college, five had completed some grad school, five were in grad school and 21 had completed grad school.

Materials

The materials consisted of the eight claims shown in Table 1 and the associated arguments, which are available in the online supplemental materials at: <https://osf.io/74pzc/>

Each argument revolved around one of the eight political claims. Each claim had six arguments associated with it, three arguments that supported the claim (*pro* arguments) and three arguments that opposed the claim (*con* arguments). For each

pro and con argument, there was a *good* version, an *inconsistent* version, and an *authority-based* version.¹

Good arguments were identical to those used in Deans-Browne and Singmann (2024). These arguments were based on arguments already established in the political discourse, and consisted of statements that strongly evidenced the arguments' conclusions.

Inconsistent arguments contained statements that provided evidence for the overall conclusion, but also contained statements that provided evidence against the overall conclusion. Like the good arguments, the statements themselves were from arguments already established in the current discourse. The evidence within these statements was also good evidence in that it strongly supported or opposed the overall conclusion of the inconsistent argument. All inconsistent arguments followed the same general structure, a statement in line with the conclusion, a statement opposing of the conclusion, and finally another statement in line with the conclusion (i.e., the inconsistent information was sandwiched between two opposing pieces of information).

Authority-based arguments emphasized the endorsement of an authority figure as evidence for its conclusion. These arguments were based on real statements celebrities, politicians, or organizations had made regarding various political issues. In this way, these arguments did not mislead participants on what authority figures had, to the best of our knowledge, actually said. The evidence in these arguments only weakly supported the argument's conclusions, in that the main evidence provided was simply that the authority figure themselves endorsed the conclusion.

Every argument was 75 words in length, and the majority of good and inconsistent arguments were inspired by arguments from www.procon.org, a non-partisan website dedicated to informing users of various political issues by providing both sides of an argument related to said issues.

Design

The Everyday Belief Bias Task consisted of two parts; initial belief ratings and subsequent argument quality ratings. Initial belief ratings were our continuous independent variable, and were on a 7-point Likert scale with anchors ranging from *extremely false* to *extremely true*. Subsequent argument quality ratings were our continuous dependent variable, and were on a 6-point Likert scale ranging from *extremely good* to *extremely bad*.

Our other independent variables related to the manipulation of the arguments themselves. One of the independent variables was *argument support*, which related to whether the argument observed was in line with the statement that had been rated (i.e., a pro argument) or was arguing against the claim that had been rated (i.e., a con argument). Participants saw exactly four pro arguments and

four con arguments each, making argument support a within-subjects design.

For both pro and con arguments, half were related to left-leaning claims (the first four claims in Table 1) and the other half were related to right-leaning claims (the last four claims in Table 1). Left-leaning and right-leaning claims were fixed so that the same claims were always left leaning and right leaning.

The other independent variable related to the manipulation of the argument was *argument type*; whether the argument presented to the participant was good, inconsistent, or authority-based. Participants always saw two pro-good and two con-good arguments, half related to statements that were left leaning and half related to statements that were right leaning. Each participant also saw one pro-inconsistent and one con-inconsistent argument each, as well as one pro-authority and one con-authority argument each. Argument type is therefore within-subjects variable. Due to limitations in the number of stimuli we had, whether each inconsistent or authority-based argument was related to a left-leaning claim or a right-leaning claim was randomly determined. The mapping of topics to the within-subject conditions was also randomly determined for each participant.

Taken together, our study implemented a 3 (argument type; good, inconsistent, or authority-based) × 2 (argument support; pro vs con) × continuous (belief; -3 to 3) within-subjects design.

Procedure

The Everyday Belief Bias Task consists of initial belief ratings and subsequent argument quality ratings. First, participants worked through each claim presented in Table 1 in a randomized order. Each claim related to a political topic that was relevant to people living in the US. For each claim, participants were tasked with rating to what extent they believe the claim.

After participants had made belief ratings for every claim, they then rated the quality of one argument for each claim, given the constraints described above. In total, half the arguments participants rated were pro arguments, and the other half were con arguments. These arguments were manipulated in terms of their informal argument quality so that half the number of arguments rated were good, a quarter of the arguments were inconsistent, and another quarter were authority based.

After participants made all argument quality ratings, they were fully debriefed. On the debrief page, participants were provided the link to unbiased arguments (mostly from www.procon.org) discussing both sides of the political claims they had seen to ensure participants did not leave the study any more ill-informed than when they entered it. The study received full ethical approval from the UCL Department of Experimental Psychology ethics committee.

¹ Due to an error in programming the experiment, one of the inconsistent arguments was incorrect, so responses to this argument were removed from analysis.

Figure 1: Argument quality ratings as a function of belief and argument type



Note. The dots show individual responses and the curved lines show predictions from the linear mixed model. Green dots represent argument quality ratings to good arguments, orange dots represent argument quality ratings to inconsistent arguments, and red dots represent argument quality ratings to authority-based arguments. The size of the dots represents the number of argument quality rating responses for the corresponding belief rating, with larger dots representing a larger number of responses. Data points are dodged so that responses for good and bad arguments do not overlap. Model predictions are based on the fixed effects of the final model that includes the non-significant three-way interaction terms. ext. = extremely.

Results

Our main research question looks to see whether we could replicate the effects seen in the Everyday Belief Bias Task with good, inconsistent, and authority-based arguments. To investigate this question, we ran a linear mixed model predicting argument quality ratings from fixed-effects of belief (with ratings ranging from -3 to 3), belief squared (consisting of the squared values of belief), argument support (pro vs con), argument type (good, inconsistent, or authority), and all interactions. The maximal model as justified by the design contained crossed random effects for the participant and argument topic grouping factors, each containing a random-effect structure mirroring the fixed-effect structure (although the participant grouping factor did not contain the highest order interaction and the argument support by argument quality interaction). This model produced a singular fit, so we simplified the model until there was no singular fit. This resulted in a final model that we use for analysis with by-participant random slopes for the main effect of argument type and by-topic random slopes for the main effects of argument support and argument type. Our process of simplifying the model, as well as the following analysis, is available in the supplemental materials: <https://osf.io/74pzc/>.

The main results of the experiment are shown in Figure 1. From this figure, we can see that participants tend to rate good arguments as having the highest quality, with inconsistent arguments being perceived as worse and

authority-based arguments being perceived worse again still. This is corroborated by the significant main effect of argument type, $F(2, 14.03) = 24.54, p < .001$. Three pairwise comparisons adjusted using the Holm method reveal that on average, good arguments are rated as significantly better than inconsistent arguments ($\text{difference} = 0.79, 95\% \text{ CI } [0.45, 1.13], t(16.53) = 6.13, p < .001$), inconsistent arguments are rated as significantly better than authority-based arguments ($\text{difference} = 0.66, 95\% \text{ CI } [0.13, 1.20], t(8.45) = 3.70, p = .006$), and good arguments are unsurprisingly rated as better than authority-based arguments ($\text{difference} = 1.45, 95\% \text{ CI } [0.90, 2.01], t(9.84) = 7.54, p < .001$).

We can also see from Figure 1 that for all types of argument, participants tend to think the quality is better when the argument is more consistent with their belief – a belief consistency effect. This is reflected in the positive linear trend for each argument type for pro arguments (which supported the claims participants rated; see Table 1), and the negative linear trend for con arguments (which challenged the claims participants rated). In line with this visual pattern, we see a belief by argument support interaction, $F(1, 593.29) = 365.04, p < .001$. For every additional point of belief, pro arguments were associated with an argument quality rating that was on average 0.30, 95% CI [0.25, 0.35], points higher ($t(909) = 12.12, p < .001$). Con arguments, in contrast, were associated with an argument quality rating that was on average 0.35, 95% CI [0.30, 0.40], points lower ($t(896) = -15.0, p < .001$).

Next, we compare the maximum belief consistency effect with the maximum effect of argument type. There are six

steps between the smallest and largest value of the belief scale, which means the effect of belief consistency from one edge of the scale to the other is roughly 0.325 (the average step change in belief ratings for pro and con arguments) multiplied by 6 which equals 1.95. This is around one third again as great as the largest argument type effect, the average difference in argument quality ratings between good and authority-based arguments, which equals 1.45.

Looking at Figure 1, the belief consistency effect does not appear to interact with the type of argument being shown, as the lines in both panels seem to be mostly parallel to each other. One potential interaction pattern would be a linear one, which would be evident in the figure if any of the lines met at one of the edges of the belief scale. If this interaction was evident in the data, we would expect to see a significant belief by argument support by argument type interaction, which is not the case, $F(2, 1092.22) = 0.43, p = .651$.

The other potential interaction pattern that could have been seen is a quadratic one, which would have been evident in the figure if any of the lines in either panel met at each edge of the belief scale. This would be evident in the data if we saw a significant squared belief by argument type interaction, which is again not the case, $F(2, 878.33) = 0.10, p = .901$. Evidence for the same interaction could also have arguably been seen by a belief squared by argument support by argument type interaction, but this was not significant either, $F(2, 1071.80) = 0.53, p = .591$.

Discussion

The goal of our research was to investigate whether people differed in their evaluations of arguments that were inconsistent compared to arguments that were based on appeals to authority. We found evidence to support this; participants on average thought that inconsistent arguments were better than authority-based arguments. However, both types of arguments were perceived as worse than good arguments. This suggests there is some aspect of authority-based arguments that is particularly unappealing which is not present in good or inconsistent arguments.

Finding that people perceive inconsistent arguments as better than authority-based arguments is somewhat surprising. In the present study, inconsistent arguments were not examples of balanced arguments like in Stanovich and West (1997) where both sides of an argument were presented as part of a dialogue. The arguments also did not provide two points of view in an organised manner. Instead, they provided conflicting evidence in place of fluent evidence, making the argument difficult to comprehend when looked at properly (see Table 2 for an example). In contrast, the authority-based arguments, did not present strong evidence for its conclusions, but did present evidence in a single direction that made it easy for the reader to parse. We speculate the preference for inconsistent over authority-based arguments stems from participants overlooking inconsistencies amongst strong evidence in favour of accepting weaker consistent evidence in an argument. Previous research suggests participants make additional inferences to ‘explain away’

inconsistencies (e.g., Otero & Kintsch, 1992; Khemlani & Johnson-Laird, 2012), though we demonstrate that such inconsistencies are not explained away completely, as participants also found inconsistent arguments less convincing than good arguments.

This has important implications. In traditional reasoning literature, the coherency of an argument is key to its quality – arguments with contradictory information are formally invalid. Whilst inconsistency as we have operationalised it is a slightly different flaw (as the inconsistent information can in theory co-exist with the consistent information), it is still surprising that participants rate arguments with such inconsistencies so highly. We are not aware of other papers investigating the perceptions of inconsistent arguments in such detail, and the present paper suggests that inconsistency within an argument may not be as important as we might assume from the literature on deductive arguments, or even as important as we think the layperson might predict.

We also replicate the belief consistency effect found previously in the Everyday Belief Bias Task (Deans-Browne & Singmann, 2024); participants think arguments are of better quality if the arguments are also in line with their beliefs. We also replicate that the maximum belief bias effect was larger than the largest effect of argument type (i.e., the average difference in argument quality ratings between good and authority-based arguments). Furthermore, we replicated the failure to find a linear or quadratic interaction between argument type and belief consistency, suggesting that the effect of argument type is somewhat similar irrespective of how in line the argument is with an individual’s belief.

We recognise the potential for task demands from eliciting participants’ prior beliefs about claims before asking them to evaluate arguments based on these claims. However we do not think this is a serious issue for the present study. We are reassured by the results from Deans-Browne and Singmann (2024), where the procedure of a similar task was made extremely transparent to participants and the order in which participants rated the claims and evaluated the arguments was manipulated. In this study, all the effects discussed in the present paper were replicated, (argument type, belief consistency, and their interaction) and were not moderated by the order manipulation.

Our results are intriguing in light of recent research by Harris et al. (2016). They argue that trustworthiness and expertise are important causal determinants when modelling the perceived convincingness of authority-based arguments in a Bayesian framework. In line with a Bayesian framework, we highlight the importance of considering prior beliefs in predicting not only the perceived quality of authority-based arguments, but also the perceived quality of good and inconsistent arguments. This leaves an open question as to what the main predictor for the perceived quality of an authority-based argument in our study is. More specifically, do participants’ prior beliefs also affect the perceived trustworthiness and expertise of the relevant authority figure in the argument, or do people with different prior beliefs agree on who is a trustworthy expert and who is not?

References

- Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, 348(6239), 1130–1132. <https://doi.org/10.1126/science.aaa1160>
- Cinelli, M., De Francisci Morales, G., Galeazzi, A., Quattrociochi, W., & Starnini, M. (2021). The echo chamber effect on social media. *Proceedings of the National Academy of Sciences*, 118(9), e2023301118. <https://doi.org/10.1073/pnas.2023301118>
- Deans-Browne, C., & Singmann, H. (2024). *The Everyday Belief Bias: Belief in the Claim of a Political Argument Predicts the Perceived Quality of the Argument*. PsyArXiv. <https://doi.org/10.31234/osf.io/x6ehu>
- Evans, J. St. B. T., Barston, J. L., & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory & Cognition*, 11(3), 295–306. <https://doi.org/10.3758/BF03196976>
- Gogarty, K., Geonzon, J., Winstanley, J., & Carter, C. (2023, October 9). Musk's X allows misinformation about Hamas' war on Israel to proliferate. *Media Matters for America*. <https://www.mediamatters.org/twitter/musks-x-allows-misinformation-about-hamas-war-israel-proliferate>
- Harris, A. J. L., Hahn, U., Madsen, J. K., & Hsu, A. S. (2016). The Appeal to Expert Opinion: Quantitative Support for a Bayesian Network Approach. *Cognitive Science*, 40(6), 1496–1533. <https://doi.org/10.1111/cogs.12276>
- Hopkins, E. J., Weisberg, D. S., & Taylor, J. C. V. (2016). The seductive allure is a reductive allure: People prefer scientific explanations that contain logically irrelevant reductive information. *Cognition*, 155, 67–76. <https://doi.org/10.1016/j.cognition.2016.06.011>
- Khemlani, S. S., & Johnson-Laird, P. N. (2012). Hidden conflicts: Explanations make inconsistencies harder to detect. *Acta Psychologica*, 139(3), 486–491. <https://doi.org/10.1016/j.actpsy.2012.01.010>
- Klauer, K. C., Musch, J., & Naumer, B. (2000). On belief bias in syllogistic reasoning. *Psychological Review*, 107(4), 852–884. <https://doi.org/10.1037/0033-295X.107.4.852>
- Menczer, F. (2021, September 20). Facebook's algorithms fueled massive foreign propaganda campaigns during the 2020 election – here's how algorithms can manipulate you. *The Conversation*. <http://theconversation.com/facebook-algorithms-fueled-massive-foreign-propaganda-campaigns-during-the-2020-election-heres-how-algorithms-can-manipulate-you-168229>
- Otero, J., & Kintsch, W. (1992). Failures to Detect Contradictions in a Text: What Readers Believe versus what they Read. *Psychological Science*, 3(4), 229–236. <https://doi.org/10.1111/j.1467-9280.1992.tb00034.x>
- Pennycook, G., & Rand, D. G. (2019). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*, 188, 39–50. <https://doi.org/10.1016/j.cognition.2018.06.011>
- Stanovich, K. E., & West, R. F. (1997). Reasoning independently of prior belief and individual differences in actively open-minded thinking. *Journal of Educational Psychology*, 89(2), 342–357. <https://doi.org/10.1037/0022-0663.89.2.342>
- Thompson, V., & Evans, J. St. B. T. (2012). Belief bias in informal reasoning. *Thinking & Reasoning*, 18(3), 278–310. <https://doi.org/10.1080/13546783.2012.670752>
- Trippas, D., Kellen, D., Singmann, H., Pennycook, G., Koehler, D. J., Fugelsang, J. A., & Dubé, C. (2018). Characterizing belief bias in syllogistic reasoning: A hierarchical Bayesian meta-analysis of ROC data. *Psychonomic Bulletin & Review*, 25(6), 2141–2174. <https://doi.org/10.3758/s13423-018-1460-7>
- Wolfe, M. B., & Kurby, C. A. (2017). Belief in the Claim of an Argument Increases Perceived Argument Soundness. *Discourse Processes*, 54(8), 599–617. <https://doi.org/10.1080/0163853X.2015.1137446>