

# Northumbria Research Link

Citation: Haigh, Matthew and Birch, Hope (2021) When 'Scientists say' coffee is good for you one day and bad for you the next: Do generic attributions to 'Scientists' and 'Experts' amplify perceived conflict? *Collabra: Psychology*, 7 (1). p. 23447. ISSN 2474-7394

Published by: University of California Press

URL: <https://doi.org/10.1525/collabra.23447> <<https://doi.org/10.1525/collabra.23447>>

This version was downloaded from Northumbria Research Link:  
<http://nrl.northumbria.ac.uk/id/eprint/46100/>

Northumbria University has developed Northumbria Research Link (NRL) to enable users to access the University's research output. Copyright © and moral rights for items on NRL are retained by the individual author(s) and/or other copyright owners. Single copies of full items can be reproduced, displayed or performed, and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided the authors, title and full bibliographic details are given, as well as a hyperlink and/or URL to the original metadata page. The content must not be changed in any way. Full items must not be sold commercially in any format or medium without formal permission of the copyright holder. The full policy is available online: <http://nrl.northumbria.ac.uk/policies.html>

This document may differ from the final, published version of the research and has been made available online in accordance with publisher policies. To read and/or cite from the published version of the research, please visit the publisher's website (a subscription may be required.)



**Northumbria  
University**  
NEWCASTLE



**UniversityLibrary**

Running Head: Scientists say...

When 'Scientists say' coffee is good for you one day and bad for you the next:  
Do generic attributions to 'Scientists' and 'Experts' amplify perceived conflict?

Matthew Haigh\*

Hope A. Birch

\*Corresponding author

Department of Psychology, Northumbria University, Newcastle upon Tyne, NE1 8ST, United Kingdom.  
Telephone: +44 (0)191 227 3472. E-mail: [matthew.haigh@northumbria.ac.uk](mailto:matthew.haigh@northumbria.ac.uk)

## Abstract

News consumers are frequently exposed to seemingly conflicting claims about the risks or benefits of activities such as eating meat and drinking coffee, which can lead to confusion and backlash against expert advice. One factor that may artificially inflate perceived conflict is the tendency for news headlines to generically attribute such claims to ‘Scientists’, ‘Experts’ or ‘Researchers’. This can create the perception that scientific consensus frequently changes, with ‘experts’ saying one thing one day (e.g., “*Fasting diet could regenerate pancreas and reverse diabetes, researchers say*”) and another the next (“*Fasting diets may raise risk of diabetes, researchers warn*”). We predicted that hedging news headlines with the qualifier ‘some’ (e.g., ...*some researchers say*) would reduce perceived contradiction and backlash by triggering the scalar inference “*some but not all...*”. We presented participants with a series of conflicting headlines or non-conflicting headlines about health and nutrition. These were presented in either their original generic format (e.g., Researchers say...) or in a qualified format (e.g., Some researchers say...). Those that saw conflicting headlines felt they were more contradictory, more confusing and resulted in us knowing less about how to be healthy than those who saw the non-conflicting headlines (Experiment 1, N=294). In Experiment 2 (N=400), the same conflict manipulation had no effect on more general beliefs about nutrition or the development of science. When our conflict manipulation did affect beliefs (Experiment 1) the effect of conflict was not moderated by headline format. Our results suggest that replacing generic consensus claims (e.g., Researchers say...) with qualified consensus claims (e.g., Some researchers say...) does not reduce the perceived contradiction and confusion that are typically associated with conflicting news reports.

**Key words:** Generics; Nutrition; Conflict; Media; Scalar Implicature; Experimental Pragmatics

News consumers are frequently exposed to seemingly conflicting claims about the risks or benefits of activities such as fasting, taking vitamin supplements, eating meat and drinking coffee. The pair of conflicting news headlines below is just one example of this phenomena.

1. Fasting diet could regenerate pancreas and reverse diabetes, researchers say  
(ABC News, February 2017)
2. Fasting diets may raise risk of diabetes, researchers warn  
(The Guardian, May 2018)

At least two-thirds of US adults report seeing or hearing media reports about nutrition at least several times a week, with over half saying that these stories conflict with earlier news reports at least some of the time (Funk & Kennedy, 2016). Understanding how people perceive such conflict has been described as a critical research need (Carpenter et al. 2016). A growing body of evidence has so far revealed that exposure to conflicting nutrition reports can create confusion (Clark et al., 2019; Nagler, 2014), negative beliefs about nutrition recommendations (Clark et al., 2019; Lee et al., 2018) and decreased engagement in food related healthy behaviour (Lee et al., 2018). Outside of the health and nutrition domain, one particularly noteworthy finding is that exposure to conflicting research studies leads many people to conclude nothing new has been learned, with some even concluding that we know *less* than before (Koehler & Pennycook, 2019).

Conflict is a normal and healthy part of the scientific process but how conflict is perceived by the public may have significant effects on how people feel towards nutritional advice and the scientific process in general. While it is important to acknowledge conflicting research findings in the media, it is also important that the reporting of science does not artificially inflate the true degree of conflict, which could lead to unnecessary confusion and backlash. In this paper we identify how media headlines that generically attribute beliefs to ‘researchers’, ‘scientists’ or ‘experts’ (such as examples 1 and 2 above) have the potential to artificially amplify perceived conflict by implying wholesale shifts in scientific consensus when conflicting research is reported (e.g., “*researchers*” *said that a fasting diet was beneficial, now they say it is harmful!*). We then go on to experimentally test whether hedging such headlines with the qualifier ‘some’ (e.g., *some researchers say...*) can moderate the sense of confusion and backlash often associated with conflicting news reports.

Consider again headlines 1 and 2, presented above. These two headlines contain information about a single behaviour (fasting) producing two distinct, directly conflicting outcomes. One headline clearly describes a benefit of a fasting diet (reversing diabetes), the other a risk (raising risk of diabetes). Previous research summarised above has shown that even brief exposure to conflicting news reports can lead to increased nutritional confusion and backlash. What these two headlines have in common is that they both generically attribute the conclusions to ‘Researchers’. These headlines make category wide assertions about the beliefs of ‘researchers’ as if researchers are a single homogeneous group that hold a high degree of consensus. In reality, each headline reports the views of a single research team based on a single study which may or may not be in line with wider scientific consensus.

In science reporting, generic claims like those above imply universal, timeless conclusions (DeJesus et al., 2019). News headlines may intentionally or unintentionally exploit the tendency for some people to

conclude that a generic statement is true for all members of the relevant category (e.g., ‘researchers say’ = ‘all researchers say’; Leslie et al., 2011; Cimpian et al., 2010). Consistent with this, Haigh, Birch & Pollet (2020) found that some participants interpreted generic phrases commonly used by the media (such as ‘Researchers say...’ or ‘Scientists believe...’) as referring to *all* relevant experts (100% of experts), while the average consensus estimate corresponded to *most* experts (i.e., >50%). This is in line with work showing that in the absence of any dissenting information the public’s default response is to assume a high degree of scientific consensus (Aklin & Urpelainen, 2014). Because there is a tendency for the public to perceive phrases such as ‘researchers believe’ as either ‘all’ or ‘most’ researchers, headlines that generically attribute claims to ‘researchers’, ‘scientists’ or ‘experts’ may amplify perceived conflict between diverging studies by implying a wholesale shift (or U-turn) in scientific consensus. In other words, they risk giving the impression that a homogenous body of experts “*say one thing one day and another the next...*” (Caswell, 2006; Goldberg & Sliwa, 2011; Jarry, 2019; Kolata, 1998; Newby, 2019; Robbins, 2012).

In the experiments that follow we examined whether introducing a simple hedge (inserting the word ‘some’ into a generic headline) could moderate the sense of conflict between diverging news stories. When directly comparing the example generic headlines above (1 and 2) there is a clear, jarring sense of contradiction (“Researchers said A now Researchers say not-A”), but this apparent U-turn may be put into perspective by inserting the qualifier ‘some’ into each headline (“Some researchers said A now Some researchers say not-A”; see examples 3 and 4).

3. Fasting diet could regenerate pancreas and reverse diabetes, *some* researchers say

4. Fasting diets may raise risk of diabetes, *some* researchers warn

The addition of qualifiers makes the diverging claims easier to reconcile, as both claims can clearly be true at the same time and do not imply a dramatic change in scientific consensus. A generic headline and the same headline qualified with ‘some’ refer to the same logical possibilities. They can both be used to describe complete consensus (i.e., it would be logically true to say ‘Some scientists think A’ when in fact ‘All scientists think A’) and they can both be used to describe less than complete consensus. However, when people encounter the word ‘some’ they typically make the pragmatic inference (known as a scalar inference or scalar implicature) that the writer or speaker is referring to ‘some but not all’ (Bott & Noveck, 2004). In the context of news headlines a phrase such as ‘Some scientists’ strongly implies the absence of complete consensus. Inserting ‘some’ as a hedge may therefore reduce the sense of contradiction by more accurately implying that some researchers believe one thing and other researchers believe another, without implying a wholesale shift in consensus.

Across two experiments we examined whether inserting the qualifier ‘some’ into genuine news headlines could reduce the consequences of perceived conflict (e.g., confusion and backlash) by making diverging claims easier to reconcile. In both experiments we presented participants with a sequence of genuine health and nutrition headlines, one at a time. They either saw a series of headlines that contained pairs of conflicting claims or a series that contained pairs of non-conflicting claims. In addition, participants saw the headlines in either their original generic format (e.g., Researchers say...) or in a qualified format with the word ‘some’ inserted (e.g., Some researchers say...). This between-subjects design meant that participants were randomly assigned to see one of four headline sequences (Generic headlines/Conflicting

claims, Generic headlines /Non-conflicting claims, Qualified headlines/Conflicting claims and Qualified headlines/Non-conflicting claims).

In Experiment 1 we predicted that participants exposed to conflicting headlines would perceive them as more contradictory than those who were exposed to non-conflicting headlines. We also predicted that relative to non-conflicting headlines, our conflicting headlines would be perceived as creating greater confusion and generating less of a cumulative increase in knowledge. Crucially, we predicted an interaction effect in which the conflicting headlines would be perceived as less contradictory and confusing when qualified with ‘some’ (with no such effect for the non-conflicting headlines).

Experiment 2 was identical in design, but with different dependent variables that measured more global beliefs about nutrition and the development of science. Previous research has shown that brief exposure to conflicting news reports is able to shift such global beliefs, at least temporarily (e.g., Clark, Nagler & Niederdeppe, 2019). We predicted that relative to non-conflicting headlines, our conflicting headlines would induce a greater sense of general Nutritional Confusion (in contrast to the more specific measure of confusion used in Experiment 1), greater Nutritional Backlash (replicating Clarke et al., 2019), greater mistrust in expertise and lower confidence in the Scientific Community. As with Experiment 1, we predicted an interaction effect in which conflicting headlines would cause less confusion, backlash and mistrust when qualified with ‘some’ (with no such effect for the non-conflicting headlines).

Experiment 2 also tested whether exposure to conflicting headlines has any positive effects, by testing the prediction that those exposed to conflicting headlines would show more sophisticated epistemic beliefs, with a greater awareness that scientific knowledge is uncertain and constantly developing (Ferguson et al., 2013; Kerwer & Rosman, 2018; Kerwer & Rosman, 2020; Kienhues et al., 2011). We predicted an effect in this positive direction as exposure to conflicting news reports may challenge erroneous perceptions that scientific knowledge is certain and unchanging. The effect was predicted to be larger when headlines were qualified with ‘some’ due to the increased emphasis on variability between individual researchers.

## **Experiments**

In two pre-registered online experiments participants were exposed to 19 genuine news headlines about human diet and nutrition. The selection of headlines seen by each participant contained either six pairs of conflicting claims (e.g., one headline reporting that alcohol is beneficial and a counterpart reporting that alcohol is not beneficial) or six pairs of non-conflicting claims (e.g., one headline reporting that dietary fat is beneficial and a counterpart also reporting that dietary fat is beneficial). Headlines were presented in either their original generic format (e.g., Scientists say...) or in a qualified format (e.g., Some scientists say...). The headlines were presented under the guise of a recognition memory task to avoid the demand characteristics that could occur if participants knew the key outcome variables related to conflict.

Both study protocols were pre-registered. All materials, raw data and code are available on the Open Science Framework (Experiment 1 <https://osf.io/eqnfg/>; Experiment 2 <https://osf.io/afrb3/>). The experiments received ethical approval through Northumbria University’s ethical approval system.

## Design

Both experiments had an identical 2x2 independent groups design. The first factor was Headline Conflict. Participants saw a series of headlines that either contained non-conflicting claims or conflicting claims about six nutrition topics. The second factor was Headline Format. The headline claims were either generically attributed to experts (e.g., “*Scientists say...*”) or were hedged using a qualifier (e.g., “*Some scientists say...*”). This resulted in four independent conditions: Generic/Conflicting, Generic/Non-conflicting, Qualified /Conflicting and Qualified /Non-conflicting. Participants were randomly assigned to one of these conditions.

In Experiment 1 the self-reported dependent variables were perceived contradiction between the headlines, level of agreement that the headlines cause confusion and a rating indicating whether the research reported in the headlines results in us knowing more or less about health and nutrition than we did before. In Experiment 2 the self-reported dependent variables were global measures of nutritional confusion, nutritional backlash, mistrust of expertise, confidence in the scientific community, beliefs about the uncertainty of scientific knowledge and beliefs about the development of scientific knowledge.

## Participants

Participants for both experiments were recruited online via the [www.prolific.co](http://www.prolific.co) participant pool which has over 70,000 registered users. The platform provides comparable data quality to the frequently used MTurk platform but with more diverse and naive users (Peer et al., 2017). Prolific report that most participants in the pool were born in the UK or USA. The majority of the pool report ethnicity as white/Caucasian and report being in either full or part-time employment. Signups are restricted based on internet protocol address and internet service provider and the number of accounts that can share the same machine are limited. Accounts cannot share PayPal or Circle accounts to avoid repeat participation. In both experiments pre-screening ensured the study was only advertised to those over 18 years old, who spoke English as their first language and had not taken part in related experiments.

In Experiment 1 we requested a sample of 300 participants (aiming for approximately 75 per group), which would give power of  $> 0.99$  to detect a medium sized interaction effect of  $f = 0.25$  (calculated using G\*Power 3.1.9.2, assuming a four groups, numerator DF = 1 and  $\alpha = 0.05$ ). A total of 312 individuals signed up through Prolific. A total of 371 survey responses were received. This discrepancy between the number of sign-ups and number of survey responses was due to a technical issue<sup>1</sup> that resulted in 59 participants participating twice (see footnote for a description of this issue). Because each participant had a unique Prolific ID we could identify those who completed the survey twice and excluded their second attempt. After doing this we were left with 312 responses from 312 participants. Following our pre-registered exclusion criteria, participants were excluded if they did not complete the task

---

<sup>1</sup> Some participants received an error message on completion of the study. Specifically, the error message appeared when they clicked a hyperlink on the debrief page. This hyperlink was inserted to redirect participants back to Prolific.co so they could claim their reward. Despite the error message, their completion of the study was successfully recorded. On seeing the error message 59 participants took it upon themselves to complete the study a second time, in the belief that their first attempt had not been recorded (we know this because some of those participants contacted us to explain their reasoning).

(considered to have withdrawn), if they declared that they did not respond seriously or if they failed an attention check (i.e., recalled < 4 headlines correctly during a headline recall task). This left a final sample of 294 participants (126 males, 168 females) which had power of 0.99 to detect a medium sized interaction effect. Participants were aged 18 - 69 ( $M_{\text{age}} = 34.29$ ,  $SD = 12.97$ ). They were paid £0.80.

In Experiment 2 We requested a sample of 400 participants (aiming for approximately 100 per group), which would give power of > 0.99 to detect a medium sized interaction effect of  $f = 0.25$ . A total of 412 individuals consented to take part. After applying our pre-registered exclusion criteria (which were the same as in Experiment 1) the final sample was 400 participants. This sample had power of >.99 to detect a medium sized interaction effect. Participants were aged 18 - 73 ( $M_{\text{age}} = 33.5$ ,  $SD = 12$ ) with 150 identifying as male, 248 identifying as female and 2 identifying with neither of those categories. They were paid £1.20.

## Materials

Both experiments were conducted online using the Qualtrics platform.

### *Stimuli*

**Headlines.** In both experiments participants saw two headlines about Vitamin D, two about red/processed meat, two about coffee, two about dietary fat, two about alcohol and two about intermittent fasting. These six topics were the focus of our study. They also saw seven unrelated filler headlines, each about a different aspect of health or nutrition.

In the Conflict conditions participants saw two headlines about each of the six topics that made conflicting claims. A pair of headlines were classed as conflicting if they implied conflicting courses of action (e.g., to drink or avoid alcohol). The following two headlines are an example of a conflicting pair.

5. Drinking wine or beer up to four times a week can protect against diabetes, researchers say (Independent, 2017)

6. Four glasses of wine is enough to harm your health, scientists say (Independent, 2014)

These specific headlines do not *directly* contradict each other (one is specific to the effects of alcohol on diabetes, the other related to more general ‘health’). However, they imply broadly conflicting conclusions about the risks of alcohol. One headline implies that the consequences of drinking alcohol are beneficial, the other implies that they are not beneficial. This type of ‘decisional’ conflict (Carpenter et al. 2016) is an important category of conflicting information to examine, as conflicting nutritional information is generally not as clear cut as in examples 1 and 2 above.

In the non-conflicting conditions, the two headlines on each topic made complimentary claims (e.g., two headlines reporting that alcohol is not beneficial, two reporting that dietary fat is beneficial, two headlines reporting that intermittent fasting is beneficial etc.). Each headline was presented on a separate page, in a fixed order. The two headlines relating to a specific topic were always presented adjacently. For the full list of items in each condition and presentation order see Appendix 1.

All headlines were sourced from Google news archive searches. We chose our six nutritional and dietary

topics as they had been the subject of conflicting news reports. We sourced the headlines by taking 24 generic phrases commonly used by the media such as ‘Scientists say’, ‘Researchers think’, and ‘Experts agree’ (these phrases were identified by Haigh et al., 2020) and pairing them with our six nutrition topics to create 144 search terms: e.g., [“scientists say”, “alcohol”]. From the search results we selected headlines that made positive or negative claims about each of our six topics. Our ‘positive’ (+) headlines either claimed a health benefit associated with an activity (e.g., *Drinking wine or beer up to four times a week can protect against diabetes, researchers say*) or the absence of presumed risk (e.g., *Scientists say eating red meat DOESN’T increase your risk of heart attack*). Our ‘negative’ (-) headlines either claimed a health risk associated with an activity (e.g., *Four glasses of wine is enough to harm your health, scientists say*) or the absence of a presumed benefit (e.g., *Taking vitamin D is pointless, say scientists*). The original headlines always made a generic claim (e.g., Scientists say...). For headlines presented in the Qualified conditions these were edited by inserting the word ‘Some’.

The seven filler headlines were sourced in a similar way by searching for nutrition and dietary topics which are often subject to media reports making claims about their consumption (e.g., organic food, turmeric). Here we used only the topic (e.g., “turmeric”) as the search term and excluded any results mentioning the authors of the scientific research and verbs that implied consensus (e.g., “Turmeric compound could boost memory and mood”). The purpose of these fillers was to disguise our manipulation. Because all participants saw the same seven filler headlines, we tested recall of these specific items in a memory recall task that served as an attention check (described below).

### Measures

**Perceived contradiction (Experiment 1).** Participants were told that of the 19 headlines they had just been asked to remember, two were about Vitamin D, two were about eating meat, two were about drinking coffee, two were about dietary fat, two were about drinking alcohol and two were about intermittent fasting. They then read six statements asking about the degree of conflict within each pair (e.g., *“The two headlines about dietary fat contradicted one another”*) and indicated the extent to which they agreed or disagreed on a 5-point Likert scale (1 = *strongly disagree*, 3 = *neither agree nor disagree*, 5 = *strongly agree*). The overall score was the sum of the six items (6 - 30), with a higher score indicating that a participant perceived a greater degree of contradiction.

**Confusion (Experiment 1).** Participants read the statement *“The headlines I was asked to remember create confusion about how to be healthy”* and rated it on a 5-point Likert scale the extent to which they agreed or disagreed (1 = *strongly disagree*, 3 = *neither agree nor disagree*, 5 = *strongly agree*).

**Scientific Advancement (Experiment 1).** Based on an item taken from Koehler & Pennycook (2019) participants were asked *“When we take the results reported in these headlines together, do we now know more, less or the same as we did before about how to be healthy?”*. They indicated their response on a 3-point scale (-1 = *we know less*, 0 = *we know the same amount*, 1 = *we know more*). This indicates whether participants perceive the body of research reported in the headlines to have increased or decreased knowledge about how to be healthy.

**Nutritional confusion (Experiment 2).** The six-item scale used by Clark et al. (2019) was used to measure confusion about nutritional advice. Participants were asked to read each statement (e.g., “I find

nutrition recommendations to be confusing”) and select on a 5-point Likert scale the point that best described them (1 = *strongly disagree*, 3 = *neither agree nor disagree*, 5 = *strongly agree*). The overall score was the average of the items (range from 1 - 5). In this study the internal reliability measured using Cronbach's alpha was 0.82.

**Nutritional backlash (Experiment 2).** The 6-item Nutritional Backlash Scale (Lee et al., 2018) was used to assess negative beliefs about nutrition recommendations and research. Participants were asked to read each statement (e.g., “Dietary recommendations are rarely useful”) and select on a 5-point Likert scale the point that best described them (1 = *strongly agree*, 3 = *neither agree nor disagree*, 5 = *strongly disagree*). The first three items were reverse scored so higher scores signified greater nutritional backlash. The overall score was the average of the items (range from 1-5). In this study the internal reliability measured using Cronbach's alpha was 0.72.

**Mistrust of expertise (Experiment 2).** We presented participants with three items from Oliver and Rahn’s (2016) populism scale that relate to mistrust of expertise. These assessed general skepticism of science and expert opinion (“I'd rather put my trust in the wisdom of ordinary people than the opinions of experts and intellectuals”, “When it comes to really important questions, scientific facts don’t help very much”, “Ordinary people can really use the help of experts to understand complicated things like science and health”). Participants read each statement and were asked to select on a 5-point Likert scale the point that best described them (1 = *strongly disagree*, 3 = *neither agree nor disagree*, 5 = *strongly agree*). The final item was reverse scored so higher scores signified greater mistrust. The overall score was the average of the items (range from 1 - 5). Cronbach's alpha was 0.73.

**Confidence in the Scientific Community (Experiment 2).** Participants were asked a question adapted from the US General Social Survey (Smith et al., 2018) which asked, “*How much confidence would you say you have in the scientific community?*” Responses were coded on a three-point scale with higher scores indicating greater mistrust (*a great deal of confidence* = 1; *only some confidence* = 2; *hardly any confidence at all* = 3).

**Epistemic beliefs about the Certainty and Development of knowledge (Experiment 2).** Two subscales from the Scientific Epistemological Beliefs Questionnaire (Conley et al., 2004) were used to measure beliefs about the certainty and development of knowledge. The first subscale had 6-items assessing beliefs about the Certainty of knowledge (belief in a right answer) e.g., “All questions in science have one right answer”. The second subscale also with 6-items assessed beliefs about the Development of scientific knowledge (beliefs about science as an evolving and changing subject) e.g., “The ideas in science books sometimes change”). Items were rated on 5-point Likert scale (1 = *strongly disagree*, 3 = *neither agree nor disagree*, 5 = *strongly agree*) with the Certainty items reverse scored. The overall score was the mean of the items (range from 1 - 5). Higher scores signified more sophisticated beliefs - i.e. a greater awareness that science is uncertain and constantly evolving. Cronbach's alpha was 0.76 for the certainty subscale and 0.82 for the development subscale.

**Headline recall test (Experiments 1 & 2).** The seven filler headlines from the list participants were asked to remember were presented alongside a further seven previously unseen headlines. Participants were told they would see a selection of seven headlines they had seen previously and seven they had not seen before. All 14 items appeared in a list and participants were asked to select the headlines they had

seen before. They could select a maximum of seven headlines. At the end of the survey participants received feedback on how many they recalled correctly. This also served as a pre-registered attention check, with participants who correctly recalled less than four headlines being excluded.

## Procedure

After reading the information sheet and providing consent, participants in both experiments were randomly assigned to see headlines in one of four headline conditions. They were presented with the list of 19 headlines making nutritional or dietary claims, with the instruction to remember these for a later recall task. The headlines were presented in the same fixed order to all participants (see Appendix 1).

Each headline was presented on its own page for a minimum of 10 seconds before a button appeared which allowed participants to move onto the next headline. There was no maximum viewing time. The headlines were numbered (e.g., Headline 1 of 19) so participants could monitor their progress. After viewing all the headlines participants in Experiment 1 answered the six conflict questions, the confusion question and scientific advancement question, before completing the recall test. Participants in Experiment 2 completed six questionnaires in the order they are listed above before completing the recall test. Mean completion time was 8.48 minutes in Experiment 1 and 10.89 minutes in Experiment 2.

In Experiment 2 each scale was presented on a separate page. Scales that required participants to select their strength of agreement were presented with “strongly disagree” on the left (1) and “strongly agree” on the right (5). The exception to this was the Nutritional Backlash scale where the order was reversed. When a participant moved on to a new page to complete the next scale, we included a clear message in bold text to indicate that the scale had changed direction (e.g., “*Note that the order of the scale has changed: Strongly agree is on the left and strongly disagree on the right*”).

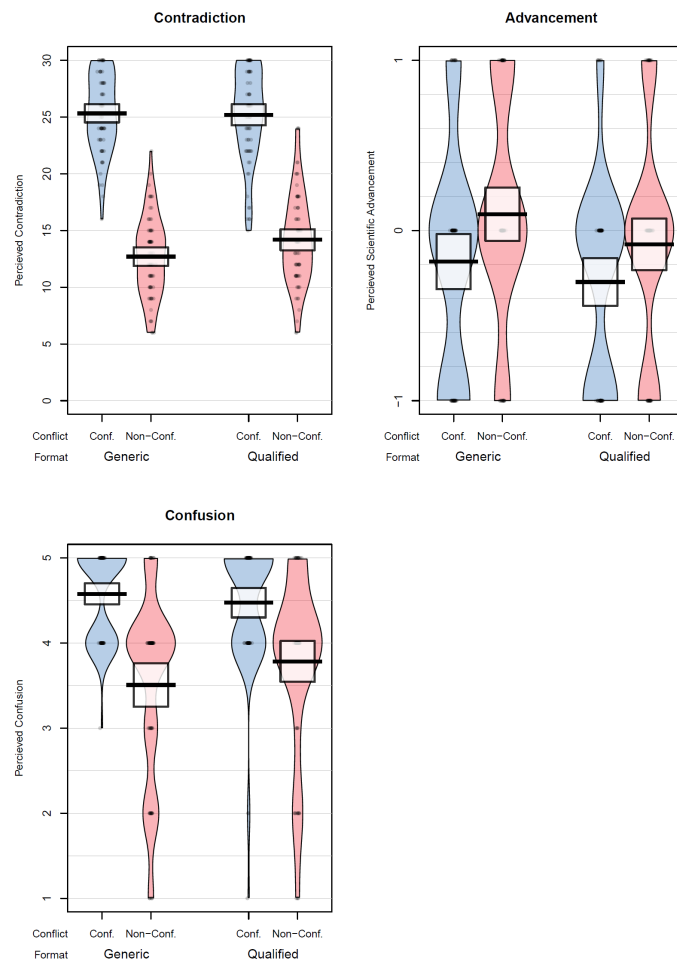
All questions required a response so participants could not progress until all items had been answered. At the end of each study, participants completed a seriousness check (“*It would be very helpful if you could tell us at this point whether you have taken part seriously, so that we can use your answers for our scientific analysis, or whether you were just clicking through to take a look at the survey?*” Aust et al., 2013) and were reassured they would be paid regardless of their response. The final page was the debrief which explained the study’s actual purpose and that the headlines should not be taken as dietary or nutrition advice. Participants' scores on the memory recall test were also provided.

## Results

Because all questions required a response there were no missing data in either experiment. A 2x2 independent groups ANOVA (with Type III sum of squares) was conducted on each DV using the car package (Fox & Weisberg, 2019) in R version 3.6.0 (R Core Team, 2019). Descriptive statistics for Experiment 1 are summarized in Figures 1 and 2 and inferential statistics are summarised in Table 1. Descriptive statistics for Experiment 2 are summarized in Figure 3 and inferential statistics are summarised in Table 2. Following the peer review process, we also present exploratory Bayesian analysis (with uninformative priors) to help interpret non-significant effects. These were conducted using the BayesFactor package in R (Morey & Rouder, 2018).

### Experiment 1

**Figure 1.** Descriptive statistics for each variable measured in Experiment 1. Black horizontal bars represent the condition mean and the band represents 95% confidence intervals



**Table 1.** ANOVA test statistics showing the effects of Conflict, Format and the interaction between Conflict and Format on each of the three dependent variables in Experiment 1.

	DF1	DF2	F	<i>p</i>	Cohen's <i>f</i> (partial)	Cohen's <i>f</i> 95% CI	BF <sub>10</sub>
<b>Contradiction</b>							
Conflict	1	290	728.28	< .001	1.58	1.41 - 1.76	>100
Format	1	290	2.43	.120	0.09	0.00 - 0.21	.21
Conflict * Format	1	290	3.41	.066	0.11	0.00 - 0.22	.16
<b>Confusion</b>							
Conflict	1	290	72.70	< .001	0.50	0.38 - 0.62	>100
Format	1	290	< 1	.402	0.05	0.00 - 0.16	.18
Conflict * Format	1	290	3.40	.066	0.11	0.00 - 0.22	.37
<b>Advancement</b>							
Conflict	1	290	10.72	.001	0.19	0.08 - 0.31	20.5
Format	1	290	3.76	.053	0.11	0.00 - 0.23	.79
Conflict * Format	1	290	< 1	.707	0.02	0.00 - 0.13	.13

### Perceived Contradiction.

*Pre-registered analysis:* The ANOVA reported in Table 1 reveals that there was a significant effect of Headline Conflict on perceived contradiction. This was measured on our contradiction scale, which had a minimum possible score of 6 and a maximum of 30. Participants exposed to conflicting headlines perceived greater contradiction between the six headline pairs ( $M = 25.3$ ) than those exposed to non-conflicting headlines ( $M = 13.4$ ). This manipulation check indicates that our conflict manipulation was perceived as intended. There was no effect of Headline Format on perceived contradiction and no interaction between Headline Conflict and Format.

*Exploratory analysis:* A Bayesian ANOVA was conducted using non-informative priors. Bayes Factors (BF<sub>10</sub>) were calculated for each effect independently (whichModels = "all"). For the Conflict manipulation our data indicate 'decisive' evidence for H<sub>1</sub> (BF<sub>10</sub> >100). For the Format manipulation our data indicate 'substantial' evidence for H<sub>0</sub> (BF<sub>10</sub> = .21). For the interaction between Conflict and Format our data also indicate 'substantial' evidence for H<sub>0</sub> (BF<sub>10</sub> = .16).

## Confusion

*Pre-registered analysis:* The ANOVA reported in Table 1 reveals that there was a significant effect of Headline Conflict on responses to our confusion item. Participants exposed to conflicting headlines indicated greater agreement that ‘*the headlines create confusion about how to be healthy*’ than those exposed to non-conflicting headlines (4.52 vs 3.65 on a 5-point scale). There was no effect of Headline Format on confusion and no interaction between Headline Conflict and Format.

*Exploratory analysis:* A Bayesian ANOVA was conducted using non-informative priors. Bayes Factors ( $BF_{10}$ ) were calculated for each effect independently (whichModels = "all"). For the Conflict manipulation our data indicate ‘decisive’ evidence for  $H_1$  ( $BF_{10} > 100$ ). For the Format manipulation our data indicate ‘substantial’ evidence for  $H_0$  ( $BF_{10} = .18$ ). For the interaction between Conflict and Format our data indicate ‘anecdotal’ evidence for  $H_0$  ( $BF_{10} = .37$ ).

## Scientific Advancement

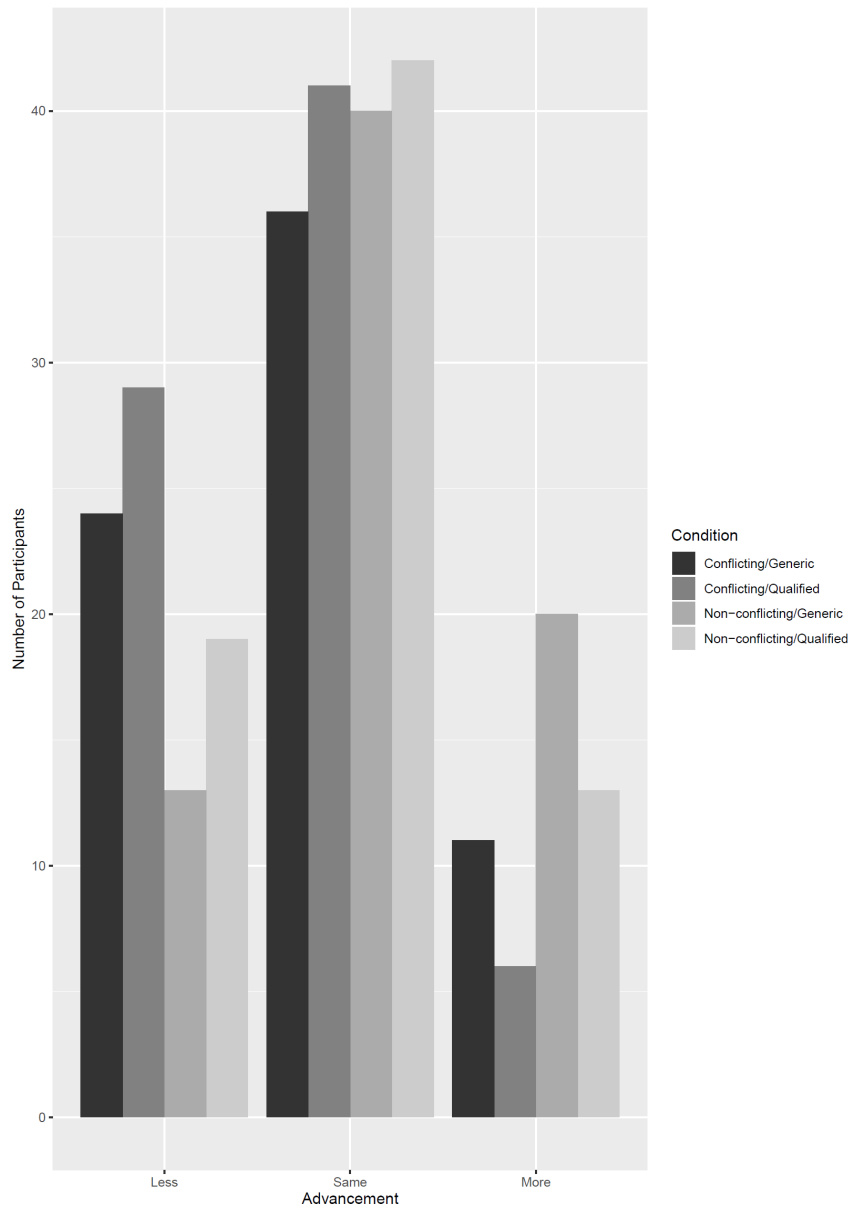
*Pre-registered analysis:* The ANOVA reported in Table 1 reveals that there was a significant effect of Headline Conflict on perceived Scientific Advancement. Advancement of knowledge was measured on a 3 point scale in which zero corresponded to ‘We know the same amount [as before]’, 1 corresponded to ‘We know more [than before]’ and -1 corresponded to ‘We know less [than before]’. The mean response of those exposed to non-conflicting headlines (0.007) was greater than the mean response of those exposed to conflicting headlines (-0.25). There was no significant effect of Headline Format on perceived Scientific Advancement and no significant interaction between Headline Conflict and Format.

*Exploratory analysis:* A Bayesian ANOVA was conducted using non-informative priors. Bayes Factors ( $BF_{10}$ ) were calculated for each effect independently (whichModels = "all"). For the Conflict manipulation our data indicate ‘strong’ evidence for  $H_1$  ( $BF_{10} = 20.5$ ). For the Format manipulation our data indicate ‘anecdotal’ evidence for  $H_0$  ( $BF_{10} = .79$ ). For the interaction between Conflict and Format our data indicate ‘substantial’ evidence for  $H_0$  ( $BF_{10} = .13$ ).

Exploratory One Sample t-tests were conducted to compare these group means to a value of zero (which is the midpoint of our scale, corresponding to ‘We know the same amount’). The mean of those exposed to non-conflicting headlines (0.007) did not significantly differ from zero ( $t(146) = 0.12, p = 0.902$ ). In contrast, the mean of those exposed to conflicting headlines (-0.25) did significantly differ from zero ( $t(146) = -4.58, p < .001$ ).

Figure 2 shows that participants exposed to conflicting headlines more frequently selected ‘we know less’, relative to those exposed to non-conflicting headlines. Mirroring this, participants exposed to conflicting headlines less frequently selected ‘we know more’. In all conditions the modal response was ‘we know the same’.

**Figure 2.** Histogram displaying the number of participants in each condition who felt the body of research reported in the headlines resulted in us knowing more than before, less than before or the same as before.

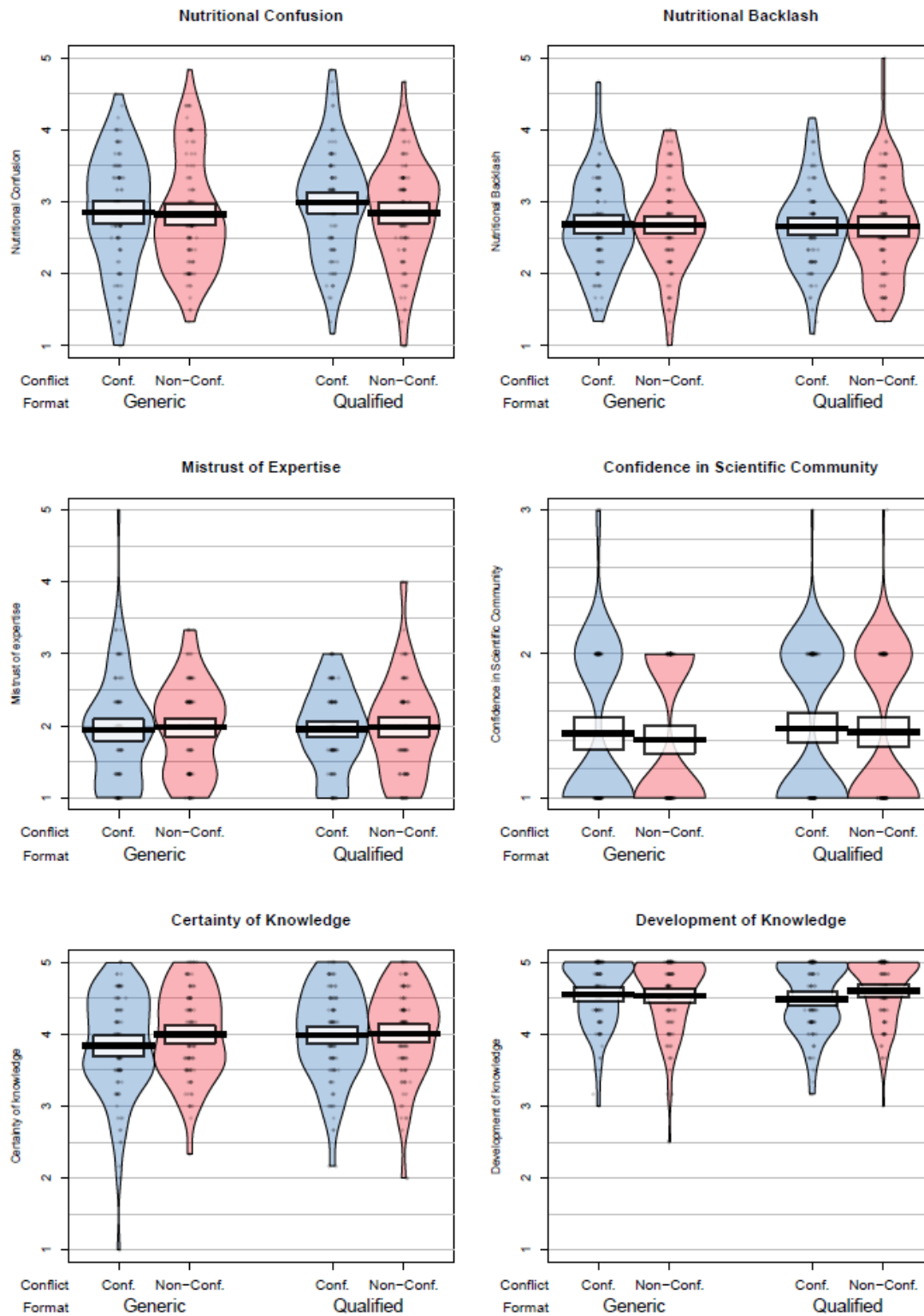


## *Experiment 2*

*Pre-registered analysis:* In Experiment 2 our pre-registered ANOVA analysis revealed that there were no significant effects of our IVs on any of the six DVs (see Table 2). The condition means are presented in Figure 3.

*Exploratory analysis:* For each of the six dependent variables in Experiment 2 a Bayesian ANOVA was conducted using non-informative priors. Using the same method as Experiment 1 Bayes Factor ( $BF_{10}$ ) were calculated for each effect independently. These are reported in Table 2. For both the Conflict and Format manipulations our data indicate ‘substantial’ evidence for  $H_0$  on all measures (all  $BF_{10} \leq .27$ ). This was also the case for the interaction between Conflict and Format with our data indicating ‘substantial’ evidence for  $H_0$  for all measures (all  $BF_{10} \leq .30$ ).

**Figure 3.** Descriptive statistics for each variable measured in Experiment 2. Black horizontal bars represent the condition mean and the band represents the 95% confidence intervals



**Table 2.** ANOVA test statistics showing the effects of Conflict, Format and the interaction between Conflict and Format on each of the six dependent variables in Experiment 2.

	DF1	DF2	F	p	Cohen's <i>f</i> (partial)	Cohen's <i>f</i> 95% CI	BF <sub>10</sub>
<b>Confusion</b>							
Conflict	1	396	1.25	.264	0.06	0.00 - 0.16	.21
Format	1	396	1.01	.315	0.06	0.00 - 0.15	.18
Conflict * Format	1	396	< 1	.458	0.04	0.00 - 0.14	.14
<b>Backlash</b>							
Conflict	1	396	< 1	.904	< 0.01	0.00 - 0.08	.11
Format	1	396	< 1	.707	0.02	0.00 - 0.11	.12
Conflict * Format	1	396	< 1	.987	< 0.01	0.00 - 0.00	.11
<b>Mistrust</b>							
Conflict	1	396	< 1	.605	0.03	0.00 - 0.12	.13
Format	1	396	< 1	.870	< 0.01	0.00 - 0.09	.11
Conflict * Format	1	396	< 1	.990	< 0.01	0.00 - 0.00	.11
<b>Confidence</b>							
Conflict	1	396	< 1	.516	0.03	0.00 - 0.13	.14
Format	1	396	< 1	.390	0.04	0.00 - 0.14	.16
Conflict * Format	1	396	< 1	.865	< 0.01	0.00 - 0.09	.11
<b>Certainty</b>							
Conflict	1	396	1.94	.165	0.07	0.00 - 0.17	.27
Format	1	396	1.60	.207	0.06	0.00 - 0.16	.23
Conflict * Format	1	396	1.02	.313	0.05	0.00 - 0.15	.18
<b>Development</b>							
Conflict	1	396	1.03	.312	0.05	0.00 - 0.15	.18
Format	1	396	< 1	.995	< 0.01	0.00 - 0.00	.11
Conflict * Format	1	396	2.06	.153	0.07	0.00 - 0.17	.30

## Discussion

Exposure to seemingly conflicting claims about the health benefits and risks of certain activities is common. This perceived conflict can lead to confusion and backlash against experts. In this paper we conducted two experiments to examine whether toning-down generic claims about expert consensus could moderate the degree of perceived conflict between diverging headlines. News headlines that generically attribute claims to ‘experts’ risk amplifying the sense of conflict by implying wholesale shifts in scientific consensus whenever the conclusions diverge from previous work. We predicted that hedging generic news headlines with a qualifier (e.g., *...some researchers say*) would reduce perceived conflict by emphasising that some experts believe one thing and other experts believe another, without implying a wholesale shift in consensus. In Experiment 1 we found that relative to those exposed to non-conflicting headlines, those exposed to conflicting headlines felt they were more contradictory, created more confusion and resulted in us knowing less about how to be healthy. Importantly, these effects were not moderated by headline Format; the effects of Conflict in the qualified condition did not differ from the effects observed in the generic condition. In Experiment 2 our Conflict manipulation did not affect more global beliefs about nutrition or the development of science, so we were unable to determine whether this effect was moderated by headline format.

In Experiment 1 we predicted that those exposed to conflicting headlines would perceive a greater sense of contradiction than those exposed to non-conflicting headlines, a greater sense of confusion and the perception that knowledge had advanced to a lesser extent. As predicted, those exposed to conflicting headlines perceived a greater sense of contradiction than those exposed to non-conflicting headlines. This manipulation check demonstrates that our conflict manipulation created the sense of conflict we intended. This was important to check as the headlines were not presented side by side and the task instructions did not tell participants to focus on the consistency of the headlines (participants were simply asked to try and remember each one). Consistent with the increased sense of contradiction, participants exposed to conflicting headlines also felt that these headlines created a greater sense of confusion about how to be healthy. Finally, participants exposed to conflicting headlines felt that the headlines advanced knowledge to a lesser extent than those exposed to non-conflicting headlines (replicating Koehler & Pennycook, 2019). Those exposed to conflicting headlines gave a significantly lower mean rating, which was also significantly lower than the midpoint of the scale; indicating that the conflicting headlines result in us knowing slightly less than we did before. This finding arguably violates normative principles of scientific inference, as new findings cannot generally reduce our knowledge (Koehler & Pennycook, 2019). A potentially useful avenue for future research would be to test whether correcting this basic misunderstanding reduces the apathy and backlash commonly associated with inconsistent findings.

Contrary to our predictions, none of the conflict effects in Experiment 1 were moderated by Headline Format (i.e., there were no significant interactions between Headline Conflict and Headline Format). The effects of our Conflict manipulation did not differ as a function of headline Format (Generic vs Qualified). In other words, adding the qualifier ‘some’ to generic headlines did not reduce the sense of contradiction and confusion associated with conflicting news reports. We predicted that inserting the word ‘some’ would trigger the scalar inference ‘some but not all’, softening the degree of perceived conflict between diverging headlines. The addition of a qualifier indicates that individual experts disagree, without implying that the same body of researchers is contradicting itself. However, this manipulation did not reduce the effects of conflict relative to those who saw the original generic headlines.

While ‘some’ has long been known to invite the scalar inference ‘some but not all’ (Bott & Noveck, 2004), one explanation for lack of interaction effect may be that generic phrases *also* invite a scalar inference. Scalar inferences are made when a speaker or writer uses a weaker quantity term (e.g., some) when a stronger term is available (e.g., all). The use of ‘some’ implies that the speaker does not believe (or does not know) that the stronger alternative holds. In other words, the writer’s choice to say ‘some scientists’ implies that the stronger term ‘all scientists’ is not believed (or not known) to be true. An explanation for our findings may be that generic phrases (e.g., Scientists say...) also invite a scalar inference. A generic phrase is not as strong as its universal equivalent (e.g., All scientists say...) and arguably not as strong as a claim indicating wide consensus (e.g., Most scientists say...). The choice to use a generic term over these stronger terms may imply that the writer does not believe or does not know whether it applies to all or even most scientists. Indeed, while previous research has shown that some people equate ‘scientists’ with ‘all scientists’ (i.e., 100% consensus) the average consensus estimate is just over half of all scientists (Haigh et al., 2020) with most participants making the ‘not all’ inference (i.e., by estimating <100% consensus). If the majority of participants made a scalar inference (i.e., not all)

from both generic and qualified headlines, it would explain why headline format did not moderate the sense of conflict. One possibility is that our headline format manipulation only buffers perceived conflict among the subset of people who equate the generic 'experts' with 'all experts'. We were unable to test this with our between-subjects design, so further research from an individual differences perspective is required to determine whether headline format moderates perceived conflict in this specific group of individuals.

In Experiment 2 we sought to test whether headline format moderated the effects of perceived conflict on more global beliefs about nutrition or the development of science. We were unable to directly test this hypothesis as our conflict manipulation (which affected perceived contradiction, confusion and advancement in Experiment 1) did not affect the more global measures used in Experiment 2. This is despite Experiment 2 having greater statistical power. We predicted that relative to non-conflicting headlines, our conflicting headlines would induce a greater sense of general Nutritional Confusion, greater Nutritional Backlash, greater mistrust in expertise and lower confidence in the Scientific Community. We also predicted that exposure to conflicting headlines may have some positive benefits, by creating greater awareness that scientific knowledge is uncertain and constantly developing. Our conflict manipulation had no effect on any of these measures, so we were unable to examine whether the expected effect was moderated by headline format. While brief exposure to conflicting headlines was sufficient to affect topic specific beliefs in Experiment 1, it did not shift the more stable and generalised beliefs measured in Experiment 2. The absence of a Conflict effect on global beliefs suggests that brief exposure to conflicting headlines was not sufficient to temporarily shift global beliefs about science or nutrition even for the short period immediately after exposure.

Hedging generic claims about scientific consensus by adding a qualifier did not affect perceived conflict in this study, however testing ways to avoid artificially inflating perceived conflict remains an important endeavour. Other methods of hedging news reports to emphasise limitations and uncertainty may be able to buffer against perceived conflict. Further research is required to examine the impact of factors such as the use of uncertainty terms in headlines (e.g., may, possibly, might), reporting of effect sizes, reporting of sample sizes and the avoidance of generics when describing different populations (e.g., sweeping claims about heterogeneous groups such as males, females and children).

## References

- ABC News. (2017, February 26). *Fasting diet could regenerate pancreas and reverse diabetes, researchers say*. <https://www.abc.net.au/news/2017-02-26/fasting-diet-could-reverse-diabetes-and-regenerate-the-pancreas/8304426>
- Aklin, M., & Urpelainen, J. (2014). Perceptions of scientific dissent undermine public support for environmental policy. *Environmental Science & Policy*, 38, 173-177. <https://doi.org/10.1016/j.envsci.2013.10.006>
- Aust, F., Diedenhofen, B., Ullrich, S., & Musch, J. (2013). Seriousness checks are useful to improve data validity in online research. *Behavior research methods*, 45(2), 527-535. <https://doi.org/10.3758/s13428-012-0265-2>
- Bohan, C. (2014, January 24). *Taking vitamin D is pointless, say scientists*. The Journal. <https://www.thejournal.ie/vitamin-d-pointless-1278578-Jan2014/>
- Boseley, S. (2014, October 7). *Studies saying fat is not that bad are misleading, scientists say*. The Guardian. <https://www.theguardian.com/society/2014/oct/07/fat-not-bad-studies-misleading-scientists-say>
- Bott, L., & Noveck, I. A. (2004). Some utterances are underinformative: The onset and time course of scalar inferences. *Journal of memory and language*, 51(3), 437-457. <https://doi.org/10.1016/j.jml.2004.05.006>
- Carbone, C. (2018, August 26). *No level of alcohol consumption is healthy, scientists say*. Fox News. <https://www.foxnews.com/science/no-level-of-alcohol-consumption-is-healthy-scientists-say>
- Carpenter, D. M., Geryk, L. L., Chen, A. T., Nagler, R. H., Dieckmann, N. F., & Han, P. K. (2016). Conflicting health information: a critical research need. *Health Expectations*, 19(6), 1173-1182. <https://doi.org/10.1111/hex.12438>
- Castrodale, J. (2019, May 14). *There is such a thing as too much coffee, scientists say*. Vice. [https://www.vice.com/en\\_us/article/evypdn/there-is-such-a-thing-as-too-much-coffee-scientists-say](https://www.vice.com/en_us/article/evypdn/there-is-such-a-thing-as-too-much-coffee-scientists-say)
- Caswell, J. A. (2006). A food scare a day: Why aren't we better at managing dietary risk?. *Human and Ecological Risk Assessment*, 12(1), 9-17. <https://doi.org/10.1080/10807030500428520>
- Cimpian, A., Brandone, A. C., & Gelman, S. A. (2010). Generic statements require little evidence for acceptance but have powerful implications. *Cognitive science*, 34(8), 1452-1482. <https://doi.org/10.1111/j.1551-6709.2010.01126.x>
- Clarke, C. E. (2008). A question of balance: The autism-vaccine controversy in the British and American elite press. *Science communication*, 30(1), 77-107. <https://doi.org/10.1177/1075547008320262>

Clark, D., Nagler, R. H., & Niederdeppe, J. (2019). Confusion and nutritional backlash from news media exposure to contradictory information about carbohydrates and dietary fats. *Public health nutrition*, 22(18), 3336-3348. <https://doi.org/10.1017/S1368980019002866>

Conley, A. M., Pintrich, P. R., Vekiri, I., & Harrison, D. (2004). Changes in epistemological beliefs in elementary science students. *Contemporary educational psychology*, 29(2), 186-204. <https://doi.org/10.1016/j.cedpsych.2004.01.004>

Day, R. (2019, June 3). *Drinking 25 cups of coffee a day is NOT bad for your heart, scientists say*. Manchester Evening News. <https://www.manchestereveningnews.co.uk/news/greater-manchester-news/drinking-25-cups-coffee-isnt-16369398>

DeJesus, J. M., Callanan, M. A., Solis, G., & Gelman, S. A. (2019). Generic language in scientific communication. *Proceedings of the National Academy of Sciences*, 116(37), 18370-18377. <https://doi.org/10.1073/pnas.1817706116>

Ferguson, L. E., Bråten, I., Strømsø, H. I., & Anmarkrud, Ø. (2013). Epistemic beliefs and comprehension in the context of reading multiple documents: Examining the role of conflict. *International Journal of Educational Research*, 62, 100-114. <https://doi.org/10.1016/j.ijer.2013.07.001>

Funk, C. & Kennedy, B. (2016) The New Food Fights: U.S. Public Divides Over Food Science. Pew Research Center.

Fox, M. (2015, September 28). *Pass the butter and pour on the olive oil, experts say*. NBC News. <https://www.nbcnews.com/health/heart-health/butter-s-not-back-pour-olive-oil-experts-say-n435166>

Fox, J. & Weisberg, S. (2019). *An R Companion to Applied Regression*. Third edition. Sage. <https://socialsciences.mcmaster.ca/jfox/Books/Companion/>.

Fukushima, K. (2014, November 9). *Ramen noodles may lead to chronic illness*. The Street. <https://www.thestreet.com/investing/stocks/ramen-noodles-may-be-bad-for-your-health-12945986>

Goldberg, J. P., & Sliwa, S. A. (2011). Communicating actionable nutrition messages: challenges and opportunities. *Proceedings of the Nutrition Society*, 70(1), 26-37. <https://doi.org/10.1017/S0029665110004714>

Hackethal, V. (2015, April 16). *Creatine linked to testicular cancer*. Web MD. <https://www.webmd.com/cancer/news/20150416/muscle-building-supplements-cancer#:~:text=April%2016%2C%202015%20%2D%2D%20Men,the%20British%20Journal%20of%20Cancer>.

Haigh, M., Birch, H. A., & Pollet, T. V. (2020). Does "Scientists believe..." imply "All scientists believe..."? Individual differences in the interpretation of generic news headlines. *Collabra: Psychology*, 6 (1), 17174. <https://doi.org/10.1525/collabra.17174>

Hannam, P. (2015, February 3). *Eating too much shark meat can damage your health*. The Sydney Morning Herald. <https://www.smh.com.au/environment/conservation/eating-too-much-shark-meat-can-damage-your-health-20150203-1351ex.html>

Hartley-Parkinson, R. (2019, October 1). *'No proof' that steak, sausages and bacon cause cancer, experts say*. Metro. <https://metro.co.uk/2019/10/01/no-proof-steak-sausages-bacon-cause-cancer-experts-say-10836881/>

Healy, M. (2015, October 26). *Hot dogs, bacon and other processed meats increase risk of cancer, scientists say*. Los Angeles Times. <https://www.latimes.com/science/la-sci-meat-dangers-20151027-story.html>

Hislop, L. (2015, March 16). *'Healthy' cereal bars contain more than 40% sugar*. The Telegraph. <https://www.telegraph.co.uk/foodanddrink/foodanddrinknews/11474537/Healthy-cereal-bars-contain-more-than-40-sugar.html#:~:text=According%20to%20consumer%20research%20group,contains%2042%20per%20cent%20sugar.>

International Food Information Council Foundation. (2018). *2018 Food and Healthy Survey*. <https://foodinsight.org/wp-content/uploads/2018/05/2018-FHS-Report-FINAL.pdf>

Jarry, J. (2019, October 11). *Why Oh Why Do Scientists Keep Changing Their Minds?* McGill University, Office for Science and Society. <https://www.mcgill.ca/oss/article/general-science-history/why-oh-why-do-scientists-keep-changing-their-minds>

Jensen, J. D., & Hurley, R. J. (2012). Conflicting stories about public scientific controversies: Effects of news convergence and divergence on scientists' credibility. *Public Understanding of Science*, 21(6), 689-704. <https://doi.org/10.1177/0963662510387759>

Jensen, J. D., Pokharel, M., Scherr, C. L., King, A. J., Brown, N., & Jones, C. (2017). Communicating uncertain science to the public: How amount and source of uncertainty impact fatalism, backlash, and overload. *Risk Analysis*, 37(1), 40-51. <https://doi.org/10.1111/risa.12600>

Kerwer, M., & Rosman, T. (2018). Mechanisms of Epistemic Change—Under Which Circumstances Does Diverging Information Support Epistemic Development?. *Frontiers in Psychology*, 9, 2278. <https://doi.org/10.3389/fpsyg.2018.02278>

Kerwer, M., & Rosman, T. (2020). Epistemic change and diverging information: How do prior epistemic beliefs affect the efficacy of short-term interventions?. *Learning and Individual Differences*, 80, 101886. <https://doi.org/10.1016/j.lindif.2020.101886>

Kienhues, D., Stadtler, M., & Bromme, R. (2011). Dealing with conflicting or consistent medical information on the web: When expert information breeds laypersons' doubts about experts. *Learning and Instruction*, 21(2), 193-204. <https://doi.org/10.1016/j.learninstruc.2010.02.004>

Knapton, S. (2016, July 21). *Everyone needs to take vitamin D supplements as modern life is bad for bones, health officials warn*. The Telegraph. <https://www.telegraph.co.uk/science/2016/07/21/everyone-needs-to-take-vitamin-d-supplements-as-modern-life-is-b/>

Koehler, D. J., & Pennycook, G. (2019). How the public, and scientists, perceive advancement of knowledge from conflicting study results. *Judgment & Decision Making*, 16(6).

Kolata, G. (1998, January 11). *The Nation: Not Quite Seconds for Everyone; The Fat's in the Fire Again*. The New York Times. <https://www.nytimes.com/1998/01/11/weekinreview/the-nation-not-quite-seconds-for-everyone-the-fat-s-in-the-fire-again.html>

Laino, C. (2011, November 30). *Eating fish may be good for brain health*. Web MD. <https://www.webmd.com/alzheimers/news/20111130/eating-fish-may-be-good-for-brain-health#1>

Leake, J. (2019, February 17). *Toast is more toxic than traffic fumes*. The Times. <https://www.thetimes.co.uk/article/toast-is-more-toxic-than-traffic-fumes-wm6pb6c8z>

Lee, C. J., Nagler, R. H., & Wang, N. (2018). Source-specific exposure to contradictory nutrition information: Documenting prevalence and effects on adverse cognitive and behavioral outcomes. *Health communication*, 33(4), 453-461. <https://doi.org/10.1080/10410236.2016.1278495>

Leslie, S. J., Khemlani, S., & Glucksberg, S. (2011). Do all ducks lay eggs? The generic overgeneralization effect. *Journal of Memory and Language*, 65(1), 15-31. <https://doi.org/10.1016/j.jml.2010.12.005>

McKie, R. (2018, May 20). *Fasting diets may raise risk of diabetes, researchers warn*. The Guardian. <https://www.theguardian.com/society/2018/may/20/fasting-diets-raise-risk-of-diabetes#:~:text=Their%20findings%20suggest%20that%20fasting,starting%20such%20programmes%20C%20say%20researchers.>

Matheson, A. (2016, December 21). *Scientists say eating red meat DOESN'T increase your risk of heart attack*. The Sun. <https://www.thesun.co.uk/living/2453093/scientists-say-eating-red-meat-doesnt-increase-your-risk-of-heart-attack/>

Molloy, A. (2014, May 17). *Four glasses of wine is enough to harm your health, scientists say*. Independent. <https://www.independent.co.uk/life-style/health-and-families/health-news/four-glasses-of-wine-is-enough-to-harm-your-health-scientists-say-9390044.html>

Morey, R., D., & Rouder, J., N. (2018). BayesFactor: Computation of Bayes Factors for Common Designs. R package version 0.9.12-4.2 <https://CRAN.R-project.org/package=BayesFactor>

Mullin, G. (2019, December 27). *Intermittent fasting diet can help you lose weight - and live longer, experts say*. The Sun. <https://www.thesun.co.uk/news/10624879/intermittent-fasting-diet-lose-weight-live-longer/>

Nagler, R. H. (2014). Adverse outcomes associated with media exposure to contradictory nutrition messages. *Journal of health communication*, 19(1), 24-40. <https://doi.org/10.1080/10810730.2013.798384>

Newby, P., K. (2019, July 13). *Feel like nutritionists are always changing their minds? Science can help you decide which diet works best for you*. Universal-Sci. <https://www.universal-sci.com/headlines/2019/7/13/feel-like-nutritionists-are-always-changing-their-minds-science-can-help-you-decide-which-diet-works-best-for-you>

Oliver, J. E., & Rahn, W. M. (2016). Rise of the Trumpenvolk: Populism in the 2016 Election. *The ANNALS of the American Academy of Political and Social Science*, 667(1), 189-206. <https://doi.org/10.1177/0002716216662639>

Paddock, C. (2015, April 17). *Maple syrup helps antibiotics defeat bacteria*. Medical News Today. <https://www.medicalnewstoday.com/articles/292540>

Park, A. (2012, March 26). *Popcorn is packed with antioxidants*. Time. <https://healthland.time.com/2012/03/26/popcorn-is-packed-with-antioxidants/>

Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, 70, 153-163. <https://doi.org/10.1016/j.jesp.2017.01.006>

Parry, L. (2015, May 7). *Always cook tomatoes and carrots but eat spinach and peppers RAW*. Daily Mail. <https://www.dailymail.co.uk/health/article-3069027/Always-cook-tomatoes-carrots-eat-spinach-peppers-RAW-Expert-reveals-nutrients-vegetables.html>

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rainey, C. (2016, February 25). *Doctors say you shouldn't start drinking coffee just because it sounds super-healthy*. New York Magazine. <https://www.grubstreet.com/2016/02/its-okay-to-not-drink-coffee.html>

Robbins, M. (2012, March 6). *"Scientists say..."*. The Guardian. <https://www.theguardian.com/science/the-lay-scientist/2012/mar/06/2>

Russell, C. (2010). Covering Controversial Science: Improving Reporting on Science and Public Policy. In D. Kennedy & G. Overholser (Eds.), *Science and the Media*. (pp 13-44). American Academy of Arts & Sciences.

Science Daily. (2019, August 5). *Rye is healthy, thanks to an interplay of microbes*. <https://www.sciencedaily.com/releases/2019/08/190805112156.htm#:~:text=Summary%3A,the%20health%20benefits%20of%20rye>.

Smith, T. W., Davern, M., Freese, J., & Morgan, S. (2018). General Social Surveys, 1972-2018. [machine-readable data file] Retrieved from <https://gssdataexplorer.norc.umd.edu/>

Spencer, B. (2018, October 4). *Vitamin D supplements do NOT make bones healthier, researchers claim*. Daily Mail. <https://www.dailymail.co.uk/health/article-6241759/NHS-advice-vitamin-D-supplements-healthier-bones-WRONG-researchers-claim.html>

Starts at 60. (2019, August 21). *Full-fat dairy and cheese good for heart health, experts now say*. <https://startsat60.com/health/health-issues/heart-health/heart-foundation-diet-recommendations-dairy-milk-cheese-eggs>

The Spokesman-Review. (2017, May 26). *Chocolate cake for breakfast is good for you*. <https://www.spokesman.com/blogs/hbo/2017/may/26/chocolate-cake-breakfast-good-you/>

Wadyka, S. (2019, February 14). *Artificially sweetened drinks linked to stroke*. Consumer Reports. <https://www.consumerreports.org/sugar-sweeteners/artificially-sweetened-drinks-linked-to-stroke/>

Whiteman, H. (2018, January 25). *Turmeric compound could boost memory and mood*. Medical News Today. <https://www.medicalnewstoday.com/articles/320732#:~:text=Researchers%20have%20found%20that%20a,and%20mood%20in%20older%20adults.>

Willey, J. (2011, June 11). *Danger in weekly fish and chips*. Express. <https://www.express.co.uk/news/uk/252029/Danger-in-weekly-fish-and-chips>

Worley, W. (2017, July 27). *Drinking wine or beer up to four times a week can protect against diabetes, researchers say*. Independent. <https://www.independent.co.uk/news/science/drinking-diabetes-wine-beer-alcohol-most-days-week-protect-insulin-denmark-a7863811.html>

## **Contributions**

Contributed to conception and design: MH, HB

Contributed to acquisition of data: MH, HB

Contributed to analysis and interpretation of data: MH, HB

Drafted and/or revised the article: MH, HB

Approved the submitted version for publication: MH, HB,

## **Funding**

This work was supported by a Leverhulme Trust Research Project Grant (RPG-2019-158) awarded to the first author.

## **Competing Interests**

The authors have no competing interests.

## **Data accessibility**

All the stimuli, presentation materials, participant data, and analysis scripts can be found on the Open Science Framework.

**Experiment 1** <https://osf.io/eqnfg/>

**Experiment 2** <https://osf.io/afrb3/>