Readers’ Regulation and Resolution of a Scientific Conflict Based on Differences in Source Information: An Eye-Tracking Study

Steffen Gottschling and Yvonne Kammerer

ABSTRACT
This eye-tracking study examines how differences in sources’ trustworthiness are used by readers to regulate and resolve conflicting scientific claims. One hundred forty-four university students were sequentially presented with two conflicting scientific claims (regarding nanotechnology) across two texts. The claims were indicated to stem from two high-trustworthiness sources, two low-trustworthiness sources, or one high-trustworthiness source and one low-trustworthiness source. After having read the claims, participants rated their subjective explanations for the conflict, their personal claim agreement, and behavioral intent and completed a source-memory task. In line with our predictions, trustworthiness differences resulted in increased visual attention to source information as compared to when both sources were of equal trustworthiness. Trustworthiness differences also affected subjective conflict explanations, claim agreement, and behavioral intent. We discuss these results in the context of the Content-Source Integration model and propose an additional differentiation between readers’ consideration of source information for conflict regulation and conflict resolution.

Introduction
Today, in the age of information, we are regularly confronted with conflicting scientific claims on unfamiliar issues, for instance, when we inform ourselves about socioscientific issues on the Internet. To illustrate this with an example, imagine that Olivia, a 22-year old psychology student who enjoys spending time outdoors, is wondering whether or not to use sunscreen containing nanoparticles. On the Internet, she encounters websites presenting the claim that sunscreen containing nanoparticles is likely to have adverse health effects. However, she also encounters other websites claiming that such sunscreen provides better protection from UV radiation and is safer than conventional products. Thus, the important question arises as to how she (as a reader with low prior domain knowledge) will deal with these conflicting scientific claims presented across multiple documents. That is, how will she try to explain and resolve the scientific conflict, and which claim (if any) might she believe in?

As Olivia (like most individuals) does not have any background in nanoscience, she could not reliably evaluate the veracity of the conflicting scientific claims based on her own personal judgment (Bromme & Goldman, 2014). However, she might be able to overcome this challenge by paying attention to and evaluating the sources that put forward the claims (“i.e., where the information comes from,” Rouet et al., 2020, p. 1) and asking herself “whom to believe” (Braasch & Scharrer, 2020; Bromme et al., 2010, 2015; Stadler & Bromme, 2014). Such strategies that involve “attending to, evaluating, and using available or accessible information about the sources of documents, such as who authored them” (Braten et al., 2017, p. 141) and that are referred to as sourcing in the research area of multiple document comprehension (e.g., Braten et al., 2017), have been shown to be positively linked...
to readers’ comprehension of information from multiple documents (e.g., Anmarkrud et al., 2014; Barzilai et al., 2015; Bråten et al., 2009; Goldman et al., 2012). While previous research has repeatedly shown a low level of sourcing in readers with low domain knowledge when reading about complex issues (Barzilai et al., 2020, 2015; Brand-Gruwel et al., 2017; Bromme et al., 2015; Gerjets et al., 2011; List & Alexander, 2019; Von Der Mühlen et al., 2016; Wiley et al., 2009; Wineburg, 1991), the presence of conflicting claims has been identified as a crucial factor to increase sourcing, as proposed by the Discrepancy-Induced Source Comprehension (D-ISC) model (Braasch & Bråten, 2017; Braasch et al., 2012). Furthermore, when encountering conflicting claims put forward by different sources, sourcing “may be triggered when the reader perceives that one source is more trustworthy than another” (Rouet et al., 2020; p. 3; also see e.g., Gottschling et al., 2019; Kammerer et al., 2016).

The goal of the present research was to explore this latter assumption in greater depth by assessing the effects of differences (as compared to no differences) in sources’ trustworthiness on readers’ visual attention to source information (by means of fine-grained eye-tracking analyses) when reading two texts that present conflicting scientific claims and their subjective explanations of the conflict, their agreement with the claims, and their resulting behavioral intents. Importantly, we aimed to not only compare a reading situation with differences in sources’ trustworthiness to a situation with two trustworthy sources but also to a situation with two untrustworthy sources, which has been widely overlooked in recent research on this topic.

**Processes of conflict evaluation and possible roles of source information**

The consideration of source information in understanding information from multiple texts is a central aspect in several theories and assumptions about multiple document comprehension, such as the Documents Model Framework (Britt et al., 1999), the Reading as Problem Solving model (Rouet et al., 2017), the D-ISC model (Braasch & Bråten, 2017; Braasch et al., 2012), or the Content-Source Integration (CSI) model (Stadtler & Bromme, 2014). While the D-ISC model contrasts reading situations with and without conflicts between sources, the CSI model focuses especially on how readers react to scientific conflicts in texts, that is, how they detect, regulate, and (potentially) resolve the conflict, and which role source information can play in readers’ conflict evaluation. Therefore, we focus on this model as a theoretical underpinning for our further considerations. The CSI model proposes that readers process scientific conflicts in three separate stages that build on each other: conflict detection, conflict regulation, and conflict resolution. During the first stage (conflict detection), contradictions between text ideas (within or between documents) are detected by the reader. This stage is a prerequisite for engaging in the subsequent stages, on which the present article focuses.

**Conflict regulation**

In the second stage, conflict regulation, readers attempt to “restore coherence after having noticed discrepant information” (Stadtler & Bromme, 2014, p. 383). According to Stadtler and Bromme, this can be achieved by ignoring the conflict, by making additional (potentially unjustified) inferences, or by acknowledging that the conflict exists due to distinct perspectives of different sources. The latter regulation strategy allows the reader to integrate the conflicting claims into a coherent mental model, as has also been proposed in the documents-as-entities assumption by Britt et al. (2013) and the D-ISC model (Braasch & Bråten, 2017). Furthermore, in some cases, source information can not only help to acknowledge the existence of a conflict but also to explain why it might have emerged (Stadtler & Bromme, 2014). Thomm and Bromme (2016) proposed that this should be the case if sources putting forward conflicting claims differ in their credibility. While perception of source credibility might be affected by various characteristics of the source, two components are considered in most conceptualizations of source credibility: expertise and trustworthiness (eg., Danielson, 2006; Hovland & Weiss, 1951; Lombardi et al., 2014; Pornpitakpan, 2004). Expertise, in this context, is the extent to which a source is perceived to be able (i.e., competent) to provide accurate and valid information. Trustworthiness is the extent to which a source is perceived to be willing to provide accurate and valid (i.e., unbiased)
information (Danielson, 2006). As proposed by Thomm and Bromme (2016), several studies indicate that when two sources that put forward conflicting scientific claims differ in these aspects of credibility, readers will be more likely to explain the conflict as due to differences in sources’ competence or sources’ motivations than when both sources seem to be equally credible (Gottschling et al., 2019, 2020; Thomm & Bromme, 2016; Thomm et al., 2015). These studies used the Explaining Conflicting Scientific Claims (ECSC; Thomm et al., 2015) questionnaire to operationalize subjective conflict explanation via agreement to statements addressing different conflict explanations. Particularly relevant for the present research are the two ECSC dimensions addressing explanations that refer to differences in characteristics of the sources, specifically, differences in researchers’ competence and researchers’ motivations. Previous research has shown that indications that two researchers putting forward conflicting claims differed in their expertise increased readers’ subjective explanations that the conflict might be due to differences in researchers’ competences (Gottschling et al., 2020; Thomm & Bromme, 2016). Likewise, indications that researchers putting forward conflicting claims differed in their trustworthiness, increased subjective explanations that the conflict might be due to differences in researchers’ motivations (Gottschling et al., 2019, 2020; Thomm & Bromme, 2016).

**Conflict resolution**

Finally, the third stage is conflict resolution, which may in part be affected by processes during the previous stage (Braasch & Scharrer, 2020). However, conflict resolution goes beyond mere explanation of a conflict. It involves the development of a personal stance toward the conflicting issue by evaluating the validity of the conflicting claims (Stadtler & Bromme, 2014). Stadtler and Bromme (2014) described two different ways of evaluating claim validity: One is a direct (first-hand) approach, in which readers evaluate the validity of the claims based on their own domain knowledge (i.e., evaluating which claim is true or accurate), whereas the other is an indirect (secondhand) approach, in which the validity of the claims is evaluated based on the perceived credibility of the sources putting forward the claims. While readers’ limited knowledge about scientific issues hinders them to use the first-hand approach reliably (Bromme & Goldman, 2014), they can still use the secondhand approach (e.g., Bromme et al., 2015) if they are able to infer the credibility of sources based on the available source information (e.g., regarding trustworthiness or expertise). This is also in line with the two-step model of validation by Richter and Maier (2017), which assumes that readers engage in more strategic validation (e.g., via source information) when initial validation processes based on knowledge activation fail to result in a sufficiently coherent representation of the issue at hand (Rouet et al., 2020).

Accordingly, in several studies, readers showed more agreement with claims from a source with high perceived trustworthiness than with claims from a source with low trustworthiness (Gottschling et al., 2019, 2020; Kammerer et al., 2016; Paul et al., 2019). Similar effects were found for claims from sources with high versus low perceived expertise (Kobayashi, 2014). Additionally, effects on the mental representation of the conflict can be found in individuals’ increased citation of sources of higher trustworthiness when writing essays about scientific topics (Braten et al., 2015; List & Alexander, 2017). In contrast, if source information does not indicate any differences in the sources’ trustworthiness or expertise, source information cannot be used to resolve contradictions between claims (Richter & Maier, 2017), which should affect the way source information is processed and used.

**Strategic processing of source information**

The D-ISc model (Braasch & Braten, 2017; Braasch et al., 2012) has proposed that source information receives increased attention in the context of conflicting as compared to consistent claims. Several studies have provided support for this assumption both in single-document (Braasch et al., 2012; Rouet et al., 2016, 2021; Saux et al., 2017, 2018, 2021) and multiple-document contexts (e.g., Kammerer et al., 2016). Since this increase in attention to source information is usually explained with strategic processing to reestablish coherence (Braasch & Braten, 2017; Braasch & Scharrer, 2020), it can be argued that source information that can be used to regulate and resolve a conflict should receive especially high attention
compared to less relevant or useful source information (Gottschling et al., 2019; Rouet et al., 2021). In line with this assumption, Saux et al. (2018) observed increased memory for source information in the presence of conflicting claims (as compared to consistent claims) when the sources differed in their indicated knowledge but not when they only differed in their described physical appearance. Further, Gottschling et al. (2019) found that differences in the trustworthiness of two sources putting forward conflicting scientific claims resulted in higher attention to source information as measured with eye-tracking (i.e., total fixation times) than when both sources were of high trustworthiness. Further, additional exploratory analyses of their eye-tracking data indicated that attention was particularly increased for the less trustworthy source when differences in source trustworthiness were present, but this was only the case for second-pass fixation durations and not for first-pass fixation durations. Second-pass fixations are generally regarded as an indication of more strategic processing, whereas first-pass fixations are considered as an indication for nonstrategic processing (Hyönä et al., 2003; Maier et al., 2018). Therefore, these findings might be additional evidence that differences in source information regarding sources’ trustworthiness are strategically processed during re-reading to explain and resolve an encountered conflict. However, as Gottschling et al. (2019) did not include a condition with two sources that both were of low trustworthiness, it cannot be excluded that increased attention to source information was actually due to the mere presence of an untrustworthy source rather than to the presence of differences in sources’ trustworthiness.

**Present study**

This study aimed to replicate and extend previous findings regarding readers’ use of source information for the regulation and resolution of an unfamiliar scientific conflict that was addressed in two texts that presented conflicting claims put forward by two different sources. Specifically, based on the findings by Gottschling et al. (2019, 2020), the main goal of the present research was to examine how source information indicating differences in sources’ trustworthiness affects readers’ regulation and resolution of the scientific conflict as compared to situations in which source information indicates that the sources are of comparable trustworthiness.

To accomplish this, in an eye-tracking experiment, participants were asked to read two conflicting scientific claims regarding the safety of two different types of nanoparticles used as UV blockers in sunscreen, with the trustworthiness of the respective sources manipulated based on three experimental conditions. Participants were explicitly made aware of the conflict between the upcoming claims during a short introduction to the topic to minimize possible effects of failure to detect the conflict on sourcing (e.g., Stang Lund et al., 2017). In the trustworthiness-differences condition, one source was an independent expert of high trustworthiness, whereas the opposing expert was potentially biased (working for a company with potential monetary interest) and therefore of lower trustworthiness. This condition was compared to two conditions in which the sources did not differ in their trustworthiness: a no-differences high-trustworthiness condition, in which both sources were indicated to be of high trustworthiness, and a no-differences low-trustworthiness condition, in which both sources were indicated to be of low trustworthiness. Thus, the main contrast to previous studies lays in the addition of a control group with two untrustworthy sources. This addition ensures that effects on increased attention to source information observed in the trustworthiness-difference condition would not only be due to the presence of an untrustworthy source but indeed due to the presence of differences in trustworthiness between the sources.

In addition to the analyses of participants’ attention to source information during reading, after participants were presented with the claims, their subjective conflict explanations, claim agreement, behavioral intent, and source memory were measured. The measure of behavioral intent was considered as a second major addition to previous research to also examine consequences of conflict resolution. Since behavioral intent is an important determinant of future behaviors (Fishbein & Ajzen, 2011), it should be incorporated more often in studies on the effects of exposure to scientific conflicts (Kobayashi, 2018) and respective sourcing processes. To this end, participants were asked about their willingness to
use products containing either one or another type of nanoparticles about which the claims were in conflict. Participants’ prior knowledge on the issue of nanoparticles in sunscreen was expected to be low (Pillai & Bezbaruah, 2017), facilitating the investigation of effects associated with the secondhand approach to evaluation (Stadtler & Bromme, 2014).

Based on the theoretical considerations and previous empirical findings outlined above, we derived the following five hypotheses, which we preregistered via aspredicted.org. Regarding conflict regulation, differences in sources’ trustworthiness should lead readers to attribute the conflict more strongly to differences in researchers’ motivations than when confronted with sources without any differences in trustworthiness (H1). Regarding conflict resolution, differences in sources’ trustworthiness should result in greater differences in agreement with the two claims (i.e., less agreement with the claim put forward by the low-trustworthiness source than with the claim of the high-trustworthiness source) than when confronted with sources without any differences in trustworthiness (H2). Likewise, differences in sources’ trustworthiness should also result in greater differences in the willingness to use the products advocated by the two sources than when confronted with sources without any differences in trustworthiness (H3).

In addition, regarding the strategic processing of source information during reading, differences in sources’ trustworthiness should result in longer total fixation duration on relevant source information than when confronted with sources without any differences in trustworthiness (H4a). More specifically, we expected this effect for second-pass fixation duration (indicating strategic processing) but not for first-pass fixation duration. However, we also expected longer total second-pass fixation duration on relevant source information when confronted with two low-trustworthiness sources than when confronted with two high-trustworthiness sources, as in Gottschling et al. (2019) when effects of differences in sources’ trustworthiness were particularly shown for the low-trustworthiness source (H4b).

Finally, due to the increased strategic processing expected in H4a, differences in sources’ trustworthiness should result in better memory for source information than when confronted with sources without any differences in trustworthiness (H5a). Furthermore, based on H4b, we also expected better source memory for the two low-trustworthiness sources than for the two high-trustworthiness sources (H5b).

Methods

Participants

Based on weighted standardized mean differences of previous studies that investigated similar effects (Gottschling et al., 2019, 2020), we expected medium effect sizes (Hedges’ g around 0.5) for group comparisons regarding the effects of differences in sources’ trustworthiness. A simulation-based a priori power analysis (conducted via bootstrapping with the intended statistical models in R (R Foundation for Statistical Computing, Vienna)) indicated a sample size of 50 participants per group to achieve a power of 80% for group comparisons, resulting in a sample size of 150 participants for our three-group design. One hundred fifty-one students from a large German university were recruited for participation. Seven participants had to be excluded from the sample because of incomplete data sets. The final sample therefore consisted of 144 university students (110 women) from different majors with a mean age of 24.18 years (SD = 4.77). The sample reported medium interest (M = 2.91, SD = 1.02) and low prior knowledge (M = 1.53, SD = 0.78) on the topic of nanotechnology, as measured with two self-reported items on a five-point scale (1 = very low to 5 = very high). All participants were compensated with 10 € for their participation in the study.

Materials

Conflict scenario and claims

The two conflicting scientific claims presented to participants were taken from the field of nanotechnology. One claim (Claim A) stated that titanium dioxide nanoparticles in sunscreen can penetrate the human skin
and therefore may cause health risks, whereas zinc oxide nanoparticles are a safe alternative. For the opposing claim (Claim B), the two types of nanoparticles were switched (for examples see Appendix A). The two texts encompassing the claims were of similar length, structure, and readability (Table 1). Before the two texts were presented, participants were given a short introduction to the topic in which they were told about the use of nanoparticles in sunscreen and the controversy about their ability to penetrate the human skin. They were told that they would subsequently be presented with two claims that were taken from the websites of two experts on the topic and that they should read them carefully to answer questions on the controversy afterward. The texts were then presented in Calibri font with 18-point font size and double-line spacing on separate HTML pages, each spanning five lines, and participants could navigate back and forth between the pages. Participants terminated the reading phase in a self-paced manner.

**Source information**

Depending on the experimental condition to which the participants were randomly assigned, different source information was added to each of the two claims. In the trustworthiness-differences condition one claim was said to stem from a professor of nanoscience working at a university who was publicly funded, whereas the other claim was said to stem from an industrially funded professor of nanoscience working for a company. In the two no-difference conditions, both sources were said to be either publicly funded professors of nanoscience working at a university (no-differences high-trustworthiness condition) or industrially funded professors of nanoscience working for a company (no-differences low-trustworthiness condition).

In the materials for the present study there were also names and countries of origin presented for both sources (Mr. Peterson from Sweden or Mr. Hendrickson from Denmark), which were identical in all conditions. For both sources it was stated that they have worked on the topic for 10 years. All source information was presented in the same location in both texts to ensure the comparability of eye-tracking data. In total, the source information for each claim encompassed additional 28 words, with 64 to 66 syllables and 195 to 197 characters, depending on the combination. The combination of sources and claims as well as the names of the sources and the order of presentation for the claims was completely counterbalanced in this study.

**Measures**

**Individual difference variables**

To ascertain comparability among our experimental conditions, we used adapted versions of the Public Knowledge of Nano Technology (PKNT) questionnaire (Lin et al., 2013) and the Public Attitudes toward Nano Technology (PANT) questionnaire (Lin et al., 2013) to measure participants’ prior knowledge on nanotechnology and their attitudes toward risks and benefits of its applications. The adapted PKNT consisted of eight multiple-choice questions about different concepts in nanotechnology (such as size and scale, structure of matter, or current applications of nanomaterials) with four response alternatives, of which only one was correct (also see Gottschling et al., 2019). The sum of correct answers was used as a measure of prior domain knowledge. For the PANT, participants were asked to rate their agreement with five statements regarding possible risks of nanotechnology for human health (e.g., “The toxicity of nanoparticles may be even higher than that of large-size particles”; Cronbach’s α = .75) and another five statements.

| Table 1. Information on Claim Material (Without Source Information) |
|-------------------------------------------------|
| Claim A | Claim B |
| Number of words | 75 | 75 |
| Number of syllables | 155 | 151 |
| Number of characters | 465 | 462 |
| Readability score^a | 58.3 | 58.3 |

^aGerman readability score (Lesbarkeitsindex; LIX)
regarding possible benefits of nanotechnology (e.g., “Nanotechnology can provide people with new and better ways to cure or examine their diseases”; Cronbach’s $\alpha = .80$) on a five-point scale ($1 = \textit{strongly disagree}$ to $5 = \textit{strongly agree}$).

**Attention to relevant source information**

Attention to source information during the reading phase was measured with eye-tracking methodology. Specifically, we measured the time (in ms) readers spent fixating on relevant source information indicating source trustworthiness (“publicly funded”/“industrially financed” and “company”/“university,” depending on the experimental group). To this end, areas of interest (AOIs) were defined around these words. Gaze points within a viewing angle of 2 degrees over a minimum period of 80 ms were defined as fixations by the SMI (SensoMotoric Instruments, Teltow) BeGaze 3.7.59 eye-tracking software used for data processing. Specifically, we computed first-pass fixation durations and second-pass fixation durations for each AOI. For the first pass, these were the sum of all fixation durations on an AOI before exiting the respective AOI; for the second pass, the sum of all fixation durations for fixations that returned to an AOI after the first pass (Hyönä et al., 2003). As the dependent variable, we used the sum of second-pass fixation durations for all AOIs aggregated across both claims. All participants had a tracking ratio of at least 80%.

**Explanation of the conflict**

After having read the conflicting claims, participants were administered the ECSC questionnaire (Thommm et al., 2015) to assess their subjective explanations for the given scientific conflict. The ECSC questionnaire uses four explanations for why researchers might contradict each other in their claims. These explanations are differences in researchers’ motivations (five items; e.g., “External factors such as competition, marketing, advertising, etc. influence the scientists in their work”; Cronbach’s $\alpha = .87$), differences in researchers’ competence (six items; e.g., “The scientists are qualified to varying degrees”; Cronbach’s $\alpha = .67$), differences in the research process (six items; e.g., “The research methods of the scientists differ, e.g., with regard to the research design or the samples”; Cronbach’s $\alpha = .81$), and thematic complexity (six items; e.g., “The topic has not yet been researched enough to be able to classify the results”; Cronbach’s $\alpha = .71$). Participants were asked to rate the extent to which each of the 23 explanatory statements might provide a potential reason for the conflict on a six-point scale ($1 = \textit{completely disagree}$ to $6 = \textit{completely agree}$).

**Claim agreement**

After having completed the ECSC questionnaire, participants were asked to rate their agreement with each of the two claims on a seven-point scale ($1 = \textit{completely disagree}$ to $7 = \textit{completely agree}$). For this purpose, the claims were presented again in the original presentation order but without source information. As the dependent variable, a difference score between the two claim agreement ratings was calculated.

**Behavioral intent**

As a measure of behavioral intent, participants were asked how likely it was that they would use sunscreen containing titanium dioxide nanoparticles and sunscreen containing zinc oxide particles, each on a seven-point scale ($1 = \textit{not at all likely}$ to $6 = \textit{very likely}$). As the dependent variable, a difference score between the two ratings was calculated.

**Source memory**

Source memory was measured with both a free- and cued-recall task for each claim. For the free-recall task, participants were given the claim and asked to give all the information they remembered regarding the source of the claim. For the cued-recall task, participants were asked directly for the name, workplace, country, and funding of the corresponding source. Separately for both tasks, source memory was scored as correct if a correct answer for the workplace (university or company) and/or
funding (publicly funded or industrially funded) was given for both sources. Otherwise, it was scored as incorrect.

**Source trustworthiness rating (manipulation check)**

At the end of the study, as a manipulation check, sources were presented again together with their respective claims and had to be rated regarding the trustworthiness of the source. Specifically, participants were asked to answer the questions “How trustworthy is this scientist in your opinion?” and “How honest is this scientist in your opinion?” on a seven-point scale (1 = not at all to 7 = very). The two items were highly correlated and thus were averaged for each source (Cronbach’s α = .91).

**Experimental design**

The study was conducted as a one-factorial between-subjects design for all dependent variables (conflict explanation, claim agreement, behavioral intent, attention to source information, and source memory). As an independent variable, differences in sources’ trustworthiness was varied between subjects with the following three conditions: a trustworthiness-differences condition, a no-differences high-trustworthiness condition, and a no-differences low-trustworthiness condition. Participants were randomly assigned to one of the three conditions, with 46 participants serving in the trustworthiness-differences condition, 50 participants in the no-differences high-trustworthiness condition, and 48 participants in the no-differences low-trustworthiness condition. To control for sequence effects of the sources in the trustworthiness-differences condition, we further divided this experimental condition into a group with the high-trustworthiness source presented first (high-low group, 25 participants) and a group with the low-trustworthiness source presented first (low-high group, 21 participants). Participants were randomly assigned to these groups.

**Procedure**

All materials were presented in German. The study was conducted in two parts. The first part was implemented as an online questionnaire and sent to the participants via the survey platform Qualtrics (Qualtrics, Provo, UT). The second part took place 24 to 48 hours later in the lab. Informed consent for participation was given at the beginning of both parts. With the online questionnaire, participants answered the items regarding prior interest and prior knowledge concerning nanotechnology and completed the PKNT and PANT questionnaires (Lin et al., 2013).

For the second part of the study in the lab, participants were given brief instructions on the eye-tracking procedure and were then seated in front of a 24-inch monitor equipped with an SMI (SensoMotoric Instruments, Teltow) RED250 mobile eye-tracking device. A chinrest ensured a constant distance of 60 cm between the monitor and the eyes of the participants. After being calibrated to the eye-tracking system (using a nine-point calibration), participants were given a short on-screen introduction to the topic and to carefully read the subsequent text materials to answer some questions about the texts afterward. Then, the two texts with the conflicting claims were presented on two separate HTML pages. Participants could navigate freely back and forth between them by clicking on respective navigation buttons, without any time constraints. To terminate the reading phase (and the eye-tracking recordings) and to proceed with the questions, participants needed to click on a respective navigation button.

After having decided to terminate the reading phase, participants were no longer able to return to the text materials. First, they were asked to complete the ECSC questionnaire (Thomm et al., 2015). Next, they had to rate their personal agreement with the two claims, which were presented in the same order as in the experimental reading phase but without any source information. Then, they were asked to rate their willingness to use sunscreen containing titanium dioxide particles and sunscreen containing zinc oxide particles. Finally, participants’ source memory was assessed separately for both claims,
also in the same order as originally presented, followed by the manipulation check to assess the perceived trustworthiness of the two sources.

**Analytic approach**

For the dependent variables conflict explanation (H1), claim agreement (H2), behavioral intent (H3), and total second-pass fixation duration on relevant source information (H4), one-factorial analyses of variance (ANOVA) with three planned contrasts were conducted. As the first contrast, we compared the trustworthiness-differences condition with the two no-differences conditions. As the second contrast, we compared the no-differences high-trustworthiness condition and the no-differences low-trustworthiness condition. If there were significant differences between the two no-differences conditions, additional simple contrasts were conducted as a control. Additionally, as a third contrast we compared the two possible sequences of source trustworthiness (high-low vs. low-high) within the trustworthiness-differences condition to control for possible sequence effects. Finally, for source memory (H5), logistic regression analyses were used (separately for the free and cued recall task) with the same contrasts as outlined above to investigate effects on readers’ probability to correctly recall both sources. All analyses in this article were conducted in R (version 4.0.3; R Core Team). All reported p-values are two-sided, and we defined statistical significance at the 5% level.

As additional exploratory analyses, linear mixed regression models with text position (first text vs. second text) as additional within-subjects predictor were conducted to further investigate effects on first-pass and second-pass fixation durations on relevant source information and the manipulation check, using the R package *nlme*. Also, mediation analyses were conducted to explore whether effects of differences in source’s trustworthiness on source memory were mediated by total second-pass fixation duration on relevant source information, using the R package *lavaan*. The dataset and all analysis scripts are available on the Open Science Framework ([https://osf.io/qt8s3/?view_only=7f5c4eb1e4f34123a0a8e46751d50720](https://osf.io/qt8s3/?view_only=7f5c4eb1e4f34123a0a8e46751d50720)).

**Results**

*Comparability of experimental conditions regarding individual difference variables*

Means (and SDs) for all individual difference variables are shown in Table 2 as a function of condition. As indicated by one-way ANOVAs with the factor *differences in sources’ trustworthiness*, experimental conditions did not differ regarding participants’ age, $F(2, 141) = 0.21, p = .808$, self-reported topic interest, $F(2, 141) = 1.25, p = .289$, self-reported prior knowledge, $F(2, 141) = 1.22, p = .298$, prior domain knowledge (PKNT score), $F(2, 141) = 0.33, p = .718$, perceived risks of nanotechnology (PANT risks score), $F(2, 141) = 0.86, p = .426$, or perceived benefits of nanotechnology (PANT benefits score), $F(2, 141) = 1.44, p = .241$. Furthermore, neither prior domain knowledge nor perceived risks and benefits of nanotechnology were significantly related to any of the dependent variables (see analysis scripts on [https://osf.io/qt8s3/?view_only=7f5c4eb1e4f34123a0a8e46751d50720](https://osf.io/qt8s3/?view_only=7f5c4eb1e4f34123a0a8e46751d50720)). Thus, we did not include them in any of the remaining analyses.

*Conflict explanation (H1)*

According to H1, we expected differences in sources’ trustworthiness to result in greater attribution of the conflict to differences in researchers’ motivations than when the sources were of comparable trustworthiness. To test this hypothesis, we first conducted a one-way multivariate ANOVA (MANOVA) for the effect of differences in source information with all four ECSC dimensions as dependent variables (i.e., differences in researchers’ motivations, differences in researchers’ competence, differences in the research process, and thematic complexity). The MANOVA showed a significant overall effect of differences in sources’ trustworthiness on conflict explanations, $F(12, 417) = 2.63, p = .002$, Pillai’s trace = 0.211, and was therefore
followed up with simple ANOVAs for the specific effects on each ECSC dimension. For the ECSC dimension “differences in researchers’ motivations,” the ANOVA showed a significant main effect of differences in sources’ trustworthiness (Figure 1), $F(3, 140) = 7.20, p < .001, \eta^2_p = .13$. In line with H1, planned contrasts revealed that in the trustworthiness-differences condition participants agreed more strongly with motivations as an explanation for the conflict ($M = 4.51, SD = 0.98$) than in the no-differences conditions ($M = 4.01, SD = 1.07$), $t(140) = 2.68, p = .008, d_{Cohen} = 0.48$. Furthermore, in the no-differences low-trustworthiness condition, participants also agreed more strongly with motivations as an explanation for the conflict ($M = 4.39, SD = 1.02$) than in the no-differences high-trustworthiness condition ($M = 3.65, SD = 1.00$), $t(140) = 3.65, p < .001, d_{Cohen} = 0.73$. An additional direct comparison between the trustworthiness-differences condition and the no-differences low-trustworthiness condition showed no significant differences for agreement with differences in researchers’ motivations as an explanation for the conflict, $t(141) = 0.59, p = .558$.

Furthermore, ANOVAs for the other dimension of the ECSC showed no significant main effects of differences in sources’ trustworthiness on the extent of attributing differences in researchers’ competence, $F(3, 140) = 1.83, p = .144$, differences in the research process, $F(3, 140) = 0.11, p = .957$, or thematic complexity, $F(3, 140) = 1.23, p = .300$, as explanations for the conflict. Means (and SDs) per condition (and trustworthiness sequence) for all four ECSC dimensions are shown in Table 2.

**Claim agreement (H2)**

According to H2, we expected greater differences in agreement with the two claims when differences in in sources’ trustworthiness were present than when the sources were of comparable trustworthiness. The ANOVA showed a significant main effect of differences in sources’ trustworthiness on the dependent variable differences in claim agreement (i.e., the difference score between the claim agreement ratings for the two claims), $F(3, 140) = 6.26, p < .001, \eta^2_p = .12$. In line with H2, planned contrasts revealed that in the trustworthiness-differences condition, the difference in claim agreement ratings was greater ($M = 1.39, SD = 1.54$) than in the two no-differences conditions ($M = 0.46, SD = 1.03$), $t(140) = 4.22, p < .001, d_{Cohen} = 0.77$, whereas the two no-differences conditions did not differ

---

**Table 2. Means (and SD) for General Measures and Dependent Variables as a Function of Differences in Sources’ Trustworthiness and Trustworthiness Sequence**

| N       | Trustworthiness Differences |           | No Differences (High) |           | No Differences (Low) |
|---------|-----------------------------|-----------|-----------------------|-----------|----------------------|
|         | High-Low                   | Low-High  | High-High             | Low-Low  |
| General measures |                       |           |                       |           |                      |
| Age, y  | 25.32 (3.99)               | 23.29 (2.81) | 24.34 (6.65)         | 23.81 (3.22) |
| Topic interest (self) | 3.08 (1.08)               | 2.67 (0.91)      | 2.76 (1.08)         | 3.08 (0.94)      |
| Topic knowledge (self) | 1.68 (0.80)               | 1.38 (0.59)      | 1.40 (0.73)         | 1.64 (0.89)      |
| PKNT    | 3.72 (1.99)               | 3.00 (1.70)      | 3.64 (1.64)         | 3.67 (1.89)      |
| PANT (risks) | 3.14 (0.79)               | 3.03 (0.62)      | 2.99 (0.57)         | 3.15 (0.57)      |
| PANT (benefits) | 3.31 (0.68)               | 3.44 (0.47)      | 3.27 (0.60)         | 3.49 (0.69)      |
| Dependent variables |                       |           |                       |           |                      |
| ECSC researchers’ competence | 3.01 (0.83)               | 2.75 (0.67)      | 2.65 (0.87)         | 2.57 (0.71)      |
| ECSC researchers’ motivations | 4.61 (1.11)               | 4.39 (0.81)      | 3.65 (1.00)         | 4.39 (1.02)      |
| ECSC research process | 4.06 (0.86)               | 4.09 (0.79)      | 4.02 (0.95)         | 3.97 (0.94)      |
| ECSC topic complexity | 4.11 (0.83)               | 3.77 (0.85)      | 4.17 (0.78)         | 4.13 (0.82)      |
| Absolute differences in claim agreement | 1.48 (1.53)               | 1.29 (1.59)      | 0.40 (0.93)         | 0.52 (1.13)      |
| Absolute differences in willingness to use | 1.12 (1.20)               | 0.76 (1.14)      | 0.44 (0.86)         | 0.58 (1.22)      |
| Total first-pass fixation duration on relevant source information, s | 3.20 (1.28)               | 2.21 (1.36)      | 2.44 (1.46)         | 2.53 (1.15)      |
| Total second-pass fixation duration on relevant source information, s | 2.13 (2.17)               | 1.72 (1.70)      | 1.04 (1.30)         | 1.19 (1.17)      |
| Source memory (free) | 48.00%                    | 52.38%       | 64.00%               | 60.41%      |
| Source memory (cued) | 56.00%                    | 61.90%       | 82.00%               | 79.17%      |
significantly, \( t(140) = 0.46, p = .625 \). The two trustworthiness sequences (high-low vs. low-high) within the trustworthiness-differences condition did also not differ significantly, \( t(140) = 0.54, p = .592 \). Means (and SDs) per condition (and trustworthiness sequence) for differences in claim agreement are shown in Table 2 and in Figure 2.

**Behavioral intent (H3)**

According to H3, we expected greater differences in the willingness to use the products advocated by the two sources when differences in sources’ trustworthiness were present than when the sources were of comparable trustworthiness. The ANOVA showed no significant main effect of differences in sources’ trustworthiness on the dependent variable differences in willingness to use the products (i.e., the difference score between the willingness to use ratings for the two types of nanoparticles), \( F(3, 140) = 2.30, p = .080, \eta_p^2 = .05 \). Still, in line with H3, planned contrasts revealed that in the trustworthiness-differences condition, the difference in participants’ willingness to use the products
ratings was greater \((M = 0.96, SD = 1.17)\) than in the two no-differences conditions \((M = 0.51, SD = 1.05)\), \(t(140) = 2.20, p = .030, d_{\text{Cohen}} = 0.41\). The two no-differences conditions did not differ significantly, \(t(140) = 0.65, p = .516\), and the two trustworthiness sequences (high-low vs. low-high) within the trustworthiness-differences condition also did not differ significantly, \(t(140) = 1.11, p = .270\). Means (and SDs) per condition (and trustworthiness sequence) for differences in willingness to use are shown in Table 2 and in Figure 3.

**Attention to source information (H4)**

According to H4a, we expected differences in sources’ trustworthiness to result in longer total second-pass fixation duration on relevant source information than when the sources were of comparable trustworthiness. Further, according to H4b, we also expected longer total second-pass fixation duration on relevant source information when confronted with two low-trustworthiness sources than when confronted with two high-trustworthiness sources. The ANOVA showed a significant main effect of differences in sources’ trustworthiness on total second-pass fixation duration on relevant source information, \(F(3, 140) = 3.54, p = .016, \eta^2_p = .07\). In line with H4a, planned contrasts revealed a longer total second-pass fixation duration (in seconds) in the trustworthiness-differences condition \((M = 1.94, SD = 1.96)\) than in the two no-differences conditions \((M = 1.11, SD = 1.23)\), \(t(140) = 3.00, p = .003, d_{\text{Cohen}} = 0.55\). However, contrary to H4b, the two no-differences conditions did not differ significantly, \(t(140) = 0.52, p = .604\), and the two trustworthiness sequences (high-low vs. low-high) within the trustworthiness-differences condition did also not differ significantly, \(t(140) = 0.93, p = .356\). Means (and SDs) per condition (and trustworthiness sequence) for total second-pass fixation duration are shown in Table 2 and in Figure 4.

To further investigate the effect of trustworthiness differences, we also computed total second-pass fixation duration on relevant source information separately for the two claims (Figure 5). In addition to the significant differences between the trustworthiness-differences conditions (high-low and low-high) and the no-trustworthiness conditions (low-low and high-high), \(t(140) = 3.01, p = .003\), a significant main effect of text position was shown, \(t(140) = 4.17, p < .001\), in that the total second-pass fixation duration on relevant source information was longer for the first text \((M = 0.88, SD = 1.06)\) than for the second text \((M = 0.49, SD = 0.76)\). There were no further significant main effects of our planned contrasts and no significant interactions, all \(p > .355\). As part of our analyses on second-pass fixation duration, we also checked readers’ navigation pattern between the texts. Overall, 21.53% of

![Figure 3](image-url). Absolute differences in willingness to use (mean) as a function of the variable differences in sources’ trustworthiness \((x\text{-axis})\) or trustworthiness sequence (shading). Error bars represent standard errors.
Readers revisited the first text after visiting the second text. This share was higher in the condition with differences in sources’ trustworthiness (32.61%; 40.00% in the high-low sequence, 23.81% in the low-high sequence) than for the conditions without differences (16.00% for high-high, 16.67% for low-low). Since without these revisits there should be no effect of differences in sources’ trustworthiness for the second-pass fixation duration on relevant source information in the first text, we conducted an additional ANOVA with our planned contrasts for this measure. In line with our assumptions, this analysis showed no significant effects, $F(3,140) = 1.12, p = .342$.

Moreover, for total first-pass fixation duration on relevant source information, the overall effect for the ANOVA with planned contrasts failed to reach significance, $F(3, 140) = 2.64, p = .052, \eta^2_p = .05$. The planned contrasts showed no significance differences in total first-pass fixation duration between the trustworthiness-differences condition and the two no-differences conditions, $t(140) = 0.92, p = .362$, and no significant differences between the two no-differences conditions, $t(140) = 0.35, p = .731$. However, the two trustworthiness sequences (high-low vs. low-high) within the trustworthiness-differences condition differed significantly, $t(140) = 2.56, p = .012, d_{Cohen} = 0.75$. Readers showed longer total first-pass fixation duration in the high-low sequence ($M = 3.20, SD = 1.28$) than in the low-high sequence ($M = 2.21, SD = 1.36$). As for the

---

**Figure 4.** Total second-pass fixation duration on relevant source information (mean) as a function of the variable differences in sources’ trustworthiness ($x$-axis) or trustworthiness sequence (shading). Error bars represent standard errors.

**Figure 5.** Total second-pass fixation duration on relevant source information (mean) as a function of trustworthiness sequence and text position. Error bars represent standard errors.
total second-pass fixation duration, we further explored this effect using linear mixed models (Figure 6). This exploratory analysis showed that the effect was driven by an interaction of trustworthiness sequence and text position in that in the high-low trustworthiness sequence the relevant source information of the second claim received significantly longer total first-pass fixation duration ($M = 1.97, SD = 0.92$) than the relevant source information of the first claim ($M = 1.23, SD = 0.65$), $t(140) = 3.61, p < .001$.

**Source memory (H5)**

According to H5, we expected differences in sources’ trustworthiness to result in better memory for source information than when the sources were of comparable trustworthiness. The logistic regression analyses for source memory showed the following results: Contrary to H5, for the free recall task, the planned contrasts showed no significance differences between the trustworthiness-differences condition and the two no-differences conditions in participants’ likelihood to correctly recall both sources, $z(140) = -1.36, p = .175$. There were also no significant differences between the two no-differences conditions, $z(140) = 0.37, p = .715$, and the two trustworthiness sequences (high-low vs. low-high) within the trustworthiness-differences condition also did not differ, $z(140) = 0.30, p = .767$.

For the cued recall task, the planned contrasts showed a significant difference between the trustworthiness-differences condition and the two no-differences conditions in participants’ likelihood to correctly recall both sources, $z(140) = -2.69, p = .007$. Contrary to H5, however, for participants in the trustworthiness-differences condition the likelihood to correctly recall both sources was lower (58.70% correct) than in the no-differences conditions (80.61% correct). There were again no significant differences between the two no-differences conditions, $z(140) = 0.35, p = .723$, and the two trustworthiness sequences (high-low vs. low-high) within the trustworthiness-differences condition also did not differ, $z(140) = 0.41, p = .686$.

**Exploratory mediation analysis for source memory**

For further exploratory analyses, we additionally conducted a mediation analysis for the differences between the trustworthiness-differences condition and the two no-differences conditions (i.e., the first contrast used in the models above) on source memory (free and cued recall) with $z$-standardized total second-pass fixation duration on source information as the mediator (Figures 7 and 8). A probit link function was used for the regression models within this analysis. The analysis follows the idea that

![Figure 6. Total first-pass fixation duration on relevant source information (mean) as a function of trustworthiness sequence and text position. Error bars represent standard errors.](image)
source memory should be affected by the visual attention allocated to source information. For the free recall task, the results showed a negative direct effect of the presence of differences in sources’ trustworthiness on the odds to recall both sources correctly, $\beta = -0.21$, 95% confidence interval (CI) = $[-0.36, -0.04]$, $p = .011$, whereas there was also an indirect positive effect via total second-pass fixation duration on source information, $\beta = 0.09$, 95% CI = $[0.02, 0.17]$, $p = .018$, as tested using a bootstrapping procedure (with 1,000 bootstrapped samples). The same analysis for the cued recall task showed similar effects with a negative direct effect, $\beta = -0.31$, 95% CI = $[-0.45, -0.14]$, $p < .001$, and positive indirect effect, $\beta = 0.09$, 95% CI = $[0.02, 0.17]$, $p = .017$. This indicates that for both the free and cued recall task the presence of differences in sources’ trustworthiness had a direct negative effect on source memory but also an indirect positive effect on source memory with total second-pass fixation duration as the mediator.

**Source trustworthiness rating (manipulation check)**

For the manipulation check we used a linear mixed model with our contrasts and text position (first text vs. second text) as additional within-subjects predictor to verify whether the source trustworthiness ratings were affected in the intended way by the manipulation of source information. On the level of experimental conditions, planned contrasts revealed significant differences between the no-differences high-trustworthiness condition and the no-differences low-trustworthiness conditions in that sources of the high-high group were rated as more trustworthy ($M = 5.38, SD = 0.93$) than the sources in the low-low group ($M = 3.76, SD = 1.39$), $t(140) = 7.53$, $p < .001$, $d_{Cohen} = 1.38$. On the level of trustworthiness sequence within the trustworthiness-differences condition, there was a significant interaction between
trustworthiness sequence and text position in that the source of the first claim was rated more trustworthy \((M = 5.74, SD = 0.90)\) than the source of the second claim \((M = 3.56, SD = 1.30)\) in the high-low group, \(t(140) = 9.43, p < .001, d_{\text{Cohen}} = 1.95\). On the contrary, the source of the second claim was rated as more trustworthy \((M = 5.21, SD = 1.02)\) than the source of the first claim in the low-high group \((M = 3.29, SD = 1.14)\), \(t(140) = 8.09, p < .001, d_{\text{Cohen}} = 1.78\). Means (and SDs) per trustworthiness sequence and text position for the manipulation check are shown in Table 3 and in Figure 9.

**Discussion**

Readers are regularly confronted with conflicting scientific claims, for instance, when they inform themselves about socioscientific issues on the Internet to make personal decisions. The CSI model by Stadtler and Bromme (2014) provides a framework for how readers choose to regulate and resolve such conflicts with the use of source information (e.g., through secondhand evaluation). The objective of this study was to expand on prior research (Gottschling et al., 2019, 2020; Thomm & Bromme, 2016) by investigating how differences in sources’ trustworthiness affect processes of conflict regulation and conflict resolution as compared to situations when sources of conflicting claims do not differ in their trustworthiness. Importantly, to the best of our knowledge, this is the first study that has compared a situation with differences in sources’ trustworthiness to both a situation with two sources of high trustworthiness and a situation with two sources of low trustworthiness. This design allowed us to examine whether increased attention to source information when differences in sources’ trustworthiness were present as shown in prior experiments (Gottschling et al., 2019) were indeed due to these differences in source information and not to the mere presence of an untrustworthy source.

**Differences in sources’ trustworthiness and conflict regulation**

Regarding our first hypothesis about readers’ conflict explanations, the present study corroborates prior research showing that source information plays a role in the regulation of scientific conflicts (Gottschling et al., 2019, 2020; Thomm & Bromme, 2016; Thomm et al., 2015). As expected in H1a, differences in sources’ trustworthiness increased participants’ attribution of the conflict to scientists’ motivations as an explanation of the conflict. This corroborates the assumption of the CSI model that one way to restore coherence is to acknowledge that the conflict is due to distinct perspectives of different sources (Stadtler & Bromme, 2014) and that source information is one potential explanation for why the conflict might have emerged (Gottschling et al., 2019; Thomm & Bromme, 2016). However, other than expected in H1b, readers’ attribution of the conflict to scientists’ motivations was higher when both sources were of low trustworthiness than when both sources were of high trustworthiness. Moreover, for those who were confronted with two low-trustworthiness sources, the extent of readers’ attribution of the conflict to scientists’ motivation was as high as for those who were confronted with differences in sources’ trustworthiness. Although we did not expect these findings, they are not surprising in hindsight: A situation with two sources of low trustworthiness might also explain the conflict in that the claims of both sources might be biased toward their vested interests. This also indicates that differences in sources’ trustworthiness are not a necessary precondition for the effect on subjective conflict explanation. We highly recommend future research to also include a condition with two low-trustworthiness sources to investigate this assumption further. Still, we also point out that source information cannot be used to resolve the conflict if both sources are of equally low trustworthiness. In other words, differences in sources’ trustworthiness might play a more essential role in conflict resolution, as we illustrate next.

**Effects on conflict resolution**

In line with our expectations regarding claim agreement, differences in sources’ trustworthiness led to increased differences in readers’ agreement with the conflicting claims compared to when both sources were either of high or low trustworthiness, corroborating the results of earlier studies (Gottschling...
Table 3. Means (and SD) for Source Trustworthiness Rating and Total Fixation Durations on Relevant Source Information as a Function of Differences in Trustworthiness Sequence and Text Position

| Text Position                        | Trustworthiness Differences |                       | No Differences (High) |                       | No Differences (Low) |                       |
|--------------------------------------|----------------------------|-----------------------|-----------------------|-----------------------|----------------------|-----------------------|
|                                      | High-Low                   | Low-High              |                       |                       |                       |                       |
| Source trustworthiness rating        | First                      | Second                | First                 | Second                | First                | Second                |
| Claim agreement                      | 5.74 (0.90)                | 3.56 (1.30)           | 3.29 (1.14)           | 5.21 (1.02)           | 5.39 (0.95)          | 5.36 (0.91)           |
| Willingness to use                   | 3.80 (1.38)                | 3.60 (1.32)           | 3.62 (1.16)           | 4.05 (1.28)           | 3.86 (1.18)          | 3.78 (1.17)           |
| Total first-pass fixation duration on | 3.16 (1.70)                | 2.60 (1.32)           | 3.19 (1.66)           | 3.19 (1.54)           | 3.30 (1.42)          | 3.14 (1.36)           |
| relevant source information, s      | 1.23 (0.65)                | 1.97 (0.92)           | 1.22 (1.04)           | 0.98 (0.61)           | 1.31 (0.90)          | 1.13 (0.84)           |
| Total second-pass fixation duration  | 1.31 (1.45)                | 0.82 (1.23)           | 0.99 (1.03)           | 0.73 (0.98)           | 0.67 (1.01)          | 0.36 (0.50)           |
| on relevant source information, s    |                            |                       |                       |                       | 0.84 (0.81)          | 0.36 (0.52)           |
Moreover, the same pattern of results was shown for differences in the willingness to use the products described in the claims as an indicator for readers’ behavioral intent. We argue that differences in claim agreement (and the differences in behavioral intent based on them) are indicators for conflict resolution as described in the CSI model (Stadtler & Bromme, 2014) as these judgments are based on the validation of the claims. In the concrete case of this study, the effects of differences in sources’ trustworthiness on claim agreement corroborate the assumption of indirect claim validation via the credibility of the respective sources. Since participants in this study had no (or only limited) prior knowledge that could be used to validate the claims directly, they turned to more strategic validation processes via the available source information (see also Bromme & Goldman, 2014; Richter & Maier, 2017). These strategic processes are also indicated by the following results regarding the attentional processes during reading.

**Additional attention allocation on relevant source information due to differences in sources’ trustworthiness**

The eye-tracking data obtained in this study provides evidence for additional allocation of visual attention on relevant source information when it indicates differences in sources’ trustworthiness as opposed to no such differences, as predicted in H4a. Specifically, as expected, this effect showed especially for total second-pass fixation duration, which are assumed to reflect strategic processing (Hyönä et al., 2003). This corroborates results of Gottschling et al. (2019), who, however, did not include a conflict situation with two low-trustworthiness sources. Other than expected in H4b, however, attention to relevant source information was not higher when the two sources were of low trustworthiness than when they were of high trustworthiness. This further supports the assumption that it is the presence of differences in sources’ trustworthiness that drives the effect of increased attention allocation to source information (i.e., to resolve the conflict) rather than the mere presence of an untrustworthy source. It should also be noted that the pattern of results for the eye-tracking data is comparable to that regarding claim agreement and behavioral intent, as indicators of conflict resolution. To summarize, this study presents further evidence that in addition to increased attention allocation to relevant source information when confronted with conflicting claims as compared to consistent claims (Braasch et al., 2012; Kammerer et al., 2016; Saux et al., 2021), particularly differences in source information that can be used to resolve a given conflict results in increased attention allocation to source information (Gottschling et al., 2019).
Finally, for source memory, the presence of differences in sources’ trustworthiness had no effect on participants’ performance in the free recall task (in line with Gottschling et al., 2019), whereas for the cued recall task differences in sources’ trustworthiness even lead to impaired (instead of enhanced) source memory compared to a situation without trustworthiness differences, thus contradicting both H5a and H5b. Furthermore, these findings contradict those of Thomm and Bromme (2016), who found better source memory when differences in source information were present. Yet, the exploratory mediation analyses conducted in this study might provide some insight into these result patterns. First, there was a direct negative effect of differences in sources’ trustworthiness on source memory for both tasks. This might be because two different source features needed to be remembered instead of twice the same. This also carried the risk to mix up these relevant source features during the recall task. In contrast, the positive indirect effects on recall performance via the attention to source information as a mediator is in line with our hypotheses. This also corroborates findings by Saux et al. (2018) showing increased source memory in the context of conflicting claims, particularly for task-relevant source information. Furthermore, according to the D-ISc model (Braasch & Bråten, 2017; Braasch et al., 2012) and the CSI model (Stadtler & Bromme, 2014), this should especially apply to source information that can be used to regulate and resolve a given conflict, given that readers are inclined to achieve coherence in their mental representation of the scientific conflict. The fact that the direct and indirect effects of differences in sources’ trustworthiness seem to work in opposite direction could explain the lack of a direct positive effect on source memory in the present study and similar experiments (Gottschling et al., 2019, 2020), despite the increased visual attention on relevant source information.

To conclude, results of the present study indicate that with differences in sources’ trustworthiness present, strategic attention allocated to relevant source information increases compared to when both sources were of high or of low trustworthiness. Further, readers seem to take the differences in sources’ trustworthiness into consideration when they rate their agreement with the two claims. Likewise, trustworthiness differences also affect readers’ behavioral intent in that they indicated they are less willing to use the product that was promoted by the low-trustworthiness source than the product that was promoted by the high-trustworthiness source. All these findings point to conflict resolution based on perceived differences in sources’ trustworthiness. In addition, our results suggest that readers also take source information into consideration in their conflict explanations, such that when at least one source is indicated to be of rather low trustworthiness they attribute the conflict more to scientists’ motivations than when both sources are indicated to be neutral and trustworthy. That is, not only when the sources differ in their trustworthiness but also when both sources are indicated to have potential vested interests (working for a nanotechnology company) do readers interpret differences in researchers’ motivations as a potential reason for the conflict. In contrast to the effects related to conflict resolution, for conflict regulation differences in sources’ trustworthiness do not seem to be a necessary precondition for source-related conflict explanation. Rather, indications that at least one source is of low trustworthiness (or biased in their motivation) seem to be sufficient. This suggests that conflict regulation and conflict resolution have different preconditions for source information to affect them. To further investigate and differentiate these preconditions seems to be promising topic for future research to understand the processes of sourcing in the evaluation of scientific conflicts. Additionally, the fact that differences in sources’ trustworthiness seem to affect claim agreement and behavioral intent (as well as conflict explanation) is in line with the assumptions of the Documents Model Framework (Britt et al., 1999) that source information is integrated into a mental representation of the conflict and, thus, can be used at a later stage to make judgments regarding the claims.

Limitations and outlook

We acknowledge that this study does not come without limitations. First, for the sake of experimental control in this study, we explicitly pointed out the presence of the scientific conflict to our participants prior to reading. This factored out the process of conflict detection and the effects it might have on sourcing during a more natural contact with scientific conflicts. In such cases validation processes (e.g.
Richter & Singer, 2017) might play an additional important role, especially when readers have more prior domain knowledge and pronounced beliefs regarding the topic. Furthermore, since this study did not focus on processes of conflict detection, we did not measure whether all readers became fully aware of the scientific conflict. Even though we tried to ensure conflict detection by pointing out the conflict in the experimental instruction, awareness of the conflict might still have varied across participants. Future studies might want to put additional focus on the detection stage of conflict evaluation. In addition, to control for individual differences in readers’ behavior, future studies could implement differences in source information as a within-subject factor instead of a between-subjects factor. However, potential carryover effects of exposure to the measures used should be kept in mind in this case. Especially questionnaires like the ECSC might prompt readers’ attention to and consideration of differences in source information after first exposure to them.

Three more points can be brought up regarding the potentially limited generalizability of our findings. First, the sample examined in this study consisted exclusively of university students, who might be more used to handling source information because of their high level of education compared to the general population. At the same time, the topic of nanoparticles in sunscreen was highly unfamiliar to them, which might have encouraged the application of a secondhand approach to evaluation (Stadtler & Bromme, 2014). Second, regarding the material of this study, we only examined one specific scientific conflict in a highly controlled experimental setting. Previous literature found that especially conflict explanation can differ based on the topic of a scientific conflict (Johnson & Dieckmann, 2018; Thomm & Bromme, 2016). Third, even though we tried to make the setting comparable to claims found on websites, actual information environments can be far more complex than the conflict presented in this study. Based on these limitations, future research should aim to replicate our findings in more natural information environments, with a more diverse sample of the population and several different conflicting topics, to test whether the effects we observed under our controlled conditions can be generalized.

Nonetheless, we believe that the findings of the present study provide intriguing new results on how readers use source information indicating low source trustworthiness or differences in sources’ trustworthiness, respectively, to regulate and resolve conflicting scientific claims about an unfamiliar socioscientific issue. Due to the large amount of conflicting scientific information and plain misinformation that can be found on the Internet and through other sources, supporting readers with low prior domain knowledge in their evaluation of scientific claims will be a challenge of ever-growing importance for education in the upcoming years.

Disclosure statement
No potential conflict of interest was reported by the author(s).

Funding
This research was in part funded by the Leibniz Research Alliance “Nanosafety.”

ORCID
Steffen Gottschling http://orcid.org/0000-0002-4458-9211
Yvonne Kammerer http://orcid.org/0000-0003-4341-517X

References
Anmarkrud, Ø., Bråten, I., & Strømso, H. I. (2014). Multiple-documents literacy: Strategic processing, source awareness, and argumentation when reading multiple conflicting documents. Learning and Individual Differences, 30, 64–76. https://doi.org/10.1016/j.lindif.2013.01.007
Barzilai, S., Thomm, E., & Shlomi-Elooz, T. (2020). Dealing with disagreement: The roles of topic familiarity and disagreement explanation in evaluation of conflicting expert claims and sources. Learning and Instruction, 69, Article 101367. https://doi.org/10.1016/j.learninstruc.2020.101367

Barzilai, S., Tzadok, E., & Eshet-Alkalai, Y. (2015). Sourcing while reading divergent expert accounts: Pathways from views of knowing to written argumentation. Instructional Science, 43(6), 737–766. https://doi.org/10.1007/s11251-015-9359-4

Braasch, J. L. G., & Bråten, I. (2017). The discrepancy-induced source comprehension (D-ISC) model: Basic assumptions and preliminary evidence. Educational Psychologist, 52(3), 167–181. https://doi.org/10.1080/00461520.2017.1323219

Braasch, J. L. G., Rouet, J.-F., Vibert, N., & Britt, M. A. (2012). Readers’ use of source information in text comprehension. Memory & Cognition, 40(3), 450–465. https://doi.org/10.3758/s13421-011-0160-6

Braasch, J. L. G., & Scharrer, L. (2020). The role of cognitive conflict in understanding and learning from multiple perspectives. In P. Van Meter, A. List, D. Lombardi, & P. Kendeou (Eds.), Handbook of learning from multiple representations and perspectives (pp. 205–222). Routledge. https://doi.org/10.4324/9780429443961-15

Brand-Gruwel, S., Kammerer, Y., Van Meeuwen, L., & Van Gog, T. (2017). Source evaluation of domain experts and novices during web search. Journal of Computer Assisted Learning, 33(3), 234–251. https://doi.org/10.1111/jcal.12162

Bråten, I., Braasch, J. L. G., Strømsø, H. I., & Ferguson, L. E. (2015). Establishing trustworthiness when students read multiple documents containing conflicting scientific evidence. Reading Psychology, 36(4), 315–349. https://doi.org/10.1080/0270711.2013.864362

Bråten, I., Stadtlater, M., & Salmerón, L. (2017). The role of sourcing in discourse comprehension. In M. F. Schober, D. N. Rapp, & A. M. Britt (Eds.), The Routledge handbook of discourse processes (2 ed, pp. 141–166). Routledge. https://doi.org/10.4324/9781315687384

Bråten, I., Stromso, H. I., & Britt, M. A. (2009). Trust matters: Examining the role of source evaluation in students’ construction of meaning within and across multiple texts. Reading Research Quarterly, 44(1), 6–28. https://doi.org/10.1598/rrq.44.1.1

Britt, M. A., Perfetti, C. A., Sandak, R., & Rouet, J.-F. (1999). Content integration and source separation in learning from multiple texts. In S. R. Goldman, A. C. Graesser, & P. Van Den Broek (Eds.), Narrative comprehension, causality, and coherence: Essays in honor of Tom Trabasso (pp. 209–233). Lawrence Erlbaum Associates.

Britt, M. A., Rouet, J.-F., & Braasch, J. L. G. (2013). Documents as entities: Extending the situation model theory of comprehension. In M. A. Britt, S. R. Goldman, & J.-F. Rouet (Eds.), Reading-from words to multiple texts (pp. 174–193). Routledge. https://doi.org/10.4324/9780203131268

Bromme, R., & Goldman, S. R. (2014). The public’s bounded understanding of science. Educational Psychologist, 49(2), 59–69. https://doi.org/10.1080/00461520.2014.921572

Bromme, R., Kienhues, D., & Porsch, T. (2010). Who knows what and who can we believe? Epistemological beliefs are beliefs about knowledge (mostly) to be attained from others. In L. D. Bendixen & F. C. Feucht (Eds.), Personal epistemology in the classroom: Theory, research, and implications for practice (pp. 163–193). Cambridge University Press. https://doi.org/10.1017/CBO9780511691904

Bromme, R., Thomm, E., & Wolf, V. (2015). From understanding to deference: Laypersons’ and medical students’ views on conflicts within medicine. International Journal of Science Education, Part B, 5(1), 68–91. https://doi.org/10.1080/21548455.2013.849017

Danielson, D. R. (2006). Web credibility. In C. Ghaoui (Ed.), Encyclopedia of human computer interaction (pp. 713–721). IGI Global. https://doi.org/10.4018/978-1-59140-562-7

Fishbein, M., & Ajzen, I. (2011). Predicting and changing behavior: The reasoned action approach. Taylor & Francis. https://doi.org/10.4324/9780203838020

Gerjets, P., Kammerer, Y., & Werner, B. (2011). Measuring spontaneous and instructed evaluation processes during web search: Integrating concurrent thinking-aloud protocols and eye-tracking data. Learning and Instruction, 21(2), 220–231. https://doi.org/10.1016/j.learninstruc.2010.02.005

Goldman, S. R., Braasch, J. L. G., Wiley, J., Graesser, A. C., & Brodowinska, K. (2012). Comprehending and learning from internet sources: Processing patterns of better and poorer learners. Reading Research Quarterly, 47(4), 356–381. https://doi.org/10.1002/RRQ.027

Gottschling, S., Kammerer, Y., & Gerjets, P. (2019). Readers’ processing and use of source information as a function of its usefulness to explain conflicting scientific claims. Discourse Processes, 56(5–6), 429–446. https://doi.org/10.1080/0163853X.2019.1610305

Gottschling, S., Kammerer, Y., Thomm, E., & Gerjets, P. (2020). How laypersons consider differences in sources’ trustworthiness and expertise in their regulation and resolution of scientific conflicts. International Journal of Science Education, Part B, 10(4), 335–354. https://doi.org/10.1080/21548455.2020.1849856

Hovland, C. I., & Weiss, W. (1951). The influence of source credibility on communication effectiveness. Public Opinion Quarterly, 15(4), 635–650. https://doi.org/10.1086/266350

Hyönen, J., Lorch, R. F., & Rinck, M. (2003). Eye movement measures to study global text processing. In J. Hyönen, R. Radach, & H. Deubel (Eds.), The mind’s eye: Cognitive and applied aspects of eye movement research (pp. 313–334). North-Holland. https://doi.org/10.1016/B978-044451020-4/50018-9
Johnson, B. B., & Dieckmann, N. F. (2018). Lay Americans’ views of why scientists disagree with each other. *Public Understanding of Science, 27*(7), 824–835. https://doi.org/10.1177/0963662517738408

Kammerer, Y., Kalbfell, E., & Gerjets, P. (2016). Is this information source commercially biased? How contradictions between web pages stimulate the consideration of source information. *Discourse Processes, 53*(5–6), 430–456. https://doi.org/10.1080/0163853X.2016.1169968

Kobayashi, K. (2014). Students’ consideration of source information during the reading of multiple texts and its effect on intertextual conflict resolution. *Instructional Science, 42*(2), 183–205. https://doi.org/10.1007/s11251-013-9276-3

Kobayashi, K. (2018). Effects of conflicting scientific arguments on belief change: Argument evaluation and expert consensus perception as mediators. *Journal of Applied Social Psychology, 48*(4), 177–187. https://doi.org/10.1111/jasp.12499

Lin, S. F., Lin, H. S., & Wu, Y. Y. (2013). Validation and exploration of instruments for assessing public knowledge of and attitudes toward nanotechnology. *Journal of Science Education and Technology, 22*(4), 548–559. https://doi.org/10.1007/s10956-012-9413-9

List, A., & Alexander, P. A. (2017). Analyzing and integrating models of multiple text comprehension. *Educational Psychologist, 52*(3), 143–147. https://doi.org/10.1080/00461520.2017.1328309

List, A., & Alexander, P. A. (2019). Toward an integrated framework of multiple text use. *In Educational Psychologist, 54*(1), 20–39. https://doi.org/10.1080/00461520.2018.1505514

Lombardi, D., Seyranian, V., & Sinatra, G. M. (2014). Source effects and plausibility judgments when reading about climate change. *Discourse Processes, 51*(1–2), 75–92. https://doi.org/10.1080/0163853X.2013.855049

Maier, J., Richter, T., & Britt, M. A. (2018). Cognitive processes underlying the text-belief consistency effect: An eye-movement study. *Applied Cognitive Psychology, 32*(2), 171–185. https://doi.org/10.1002/acp.3391

Paul, J., Stadtlper, M., & Bromme, R. (2019). Effects of a sourcing prompt and conflicts in reading materials on elementary students’ use of source information. *Discourse Processes, 56*(2), 155–169. https://doi.org/10.1080/0163853X.2017.1402165

Pillai, R. G., & Bezbahna, A. N. (2017). Perceptions and attitude effects on nanotechnology acceptance: An exploratory framework. *Journal of Nanoparticle Research, 19*(2), 41–54. https://doi.org/10.1007/s11051-016-3733-2

Pornpitakpan, C. (2004). The persuasiveness of source credibility: A critical review of five decades’ evidence. *Journal of Applied Social Psychology, 34*(2), 243–281. https://doi.org/10.1111/j.1559-1816.2004.tb02547.x

Richter, T., & Maier, J. (2017). Comprehension of multiple documents with conflicting information: A two-step model of validation. *Educational Psychologist, 52*(3), 148–166. https://doi.org/10.1080/00461520.2017.1322968

Richter, T., & Singer, M. (2017). Discourse updating: Acquiring and revising knowledge through discourse. In M. F. Schober, D. N. Rapp & M. A. Britt (Eds.), *The Routledge handbook of discourse processes* (pp. 167–190). Routledge.

Rouet, J.-F., Britt, M. A., & Durik, A. M. (2017). RESOLV: Readers’ representation of reading contexts and tasks. *Educational Psychologist, 52*(3), 200–215. https://doi.org/10.1080/00461520.2017.1329015

Rouet, J.-F., Le Bigot, L., De Pereyra, G., & Britt, M. A. (2016). Whose story is this? Discrepancy triggers readers’ attention to source information in short narratives. *Reading and Writing, 29*(8), 1549–1570. https://doi.org/10.1007/s11145-016-9625-0

Rouet, J.-F., Ros, C., Bordas, B., Sanchiz, M., Saux, G., Richter, T., & Britt, M. A. (2020). *When does source information help? Content vs. source-based validation as a function of readers’ prior knowledge* (No. 4493; EasyChair Preprint), EasyChair. https://easychair.org/publications/preprint/2w2N

Rouet, J.-F., Saux, G., Ros, C., Stadtlper, M., Vibert, N., & Britt, M. A. (2021). Inside document models: Role of source attributes in readers’ integration of multiple text contents. *Discourse Processes, 58*(1), 60–79. https://doi.org/10.1080/0163853X.2020.1750246

Saux, G., Britt, A., Le Bigot, L., Vibert, N., Burin, D., & Rouet, J. F. (2017). Conflicting but close: Readers’ integration of information sources as a function of their disagreement. *Memory and Cognition, 45*(1), 151–167. https://doi.org/10.3758/s13421-016-0644-5

Saux, G., Ros, C., Britt, M. A., Stadtlper, M., Burin, D. I., & Rouet, J. F. (2018). Readers’ selective recall of source features as a function of claim discrepancy and task demands. *Discourse Processes, 55*(5–6), 525–544. https://doi.org/10.1080/0163853X.2018.1463722

Saux, G., Vibert, N., Dampuré, J., Burin, D. I., Britt, M. A., & Rouet, J. F. (2021). From simple agents to information sources: Readers’ differential processing of story characters as a function of story consistency. *Acta Psychologica, 212* (103191), 1–16. https://doi.org/10.1016/j.actpsy.2020.103191

Stadtlper, M., & Bromme, R. (2014). The content–source integration model: A taxonomic description of how readers comprehend conflicting scientific information. In D. N. Rapp & J. L. G. Braasch (Eds.), *Processing inaccurate information: Theoretical and applied perspectives from cognitive science and the educational sciences* (pp. 379–402). MIT Press.

Stang Lund, E., Bråten, I., Brante, E. W., & Strømso, H. I. (2017). Memory for textual conflicts predicts sourcing when adolescents read multiple expository texts. *Reading Psychology, 38*(4), 417–437. https://doi.org/10.1080/02702711.2016.1278417
Thomm, E., & Bromme, R. (2016). How source information shapes lay interpretations of science conflicts: Interplay between sourcing, conflict explanation, source evaluation, and claim evaluation. Reading and Writing, 29, 1629–1652. https://doi.org/10.1007/s11145-016-9638-8

Thomm, E., Hentschke, J., & Bromme, R. (2015). The explaining conflicting scientific claims (ECSC) questionnaire: Measuring laypersons' explanations for conflicts in science. Learning and Individual Differences, 37, 139–152. https://doi.org/10.1016/j.lindif.2014.12.001

Von Der Mühlen, S., Richter, T., Schmid, S., Schmidt, E. M., & Berthold, K. (2016). The use of source-related strategies in evaluating multiple psychology texts: A student–scientist comparison. Reading and Writing, 29(8), 1677–1698. https://doi.org/10.1007/s11145-015-9601-0

Wiley, J., Goldman, S. R., Graesser, A. C., Sanchez, C. A., Ash, I. K., & Hemmerich, J. A. (2009). Source evaluation, comprehension, and learning in internet science inquiry tasks. American Educational Research Journal, 46(4), 1060–1106. https://doi.org/10.3102/0002831209333183

Wineburg, S. S. (1991). Historical problem solving: A study of the cognitive processes used in the evaluation of documentary and pictorial evidence. Journal of Educational Psychology, 83(1), 73–87. https://doi.org/10.1037/0022-0663.83.1.73

Appendix A

Two versions of the claims used in the study (translated from German) are presented. Relevant source information that was encompassed by AOIs for eye-tracking analysis is printed in bold.

**Claim in favor of zinc oxide**

Mr. Hendricksen, a **publicly funded** professor working in nanoscience at a Danish **university**, believes that titanium dioxide nanoparticles in particular can be expected to penetrate the upper layers of the skin and therefore have an undesirable effect on our health. The scientist has been working on this topic at his **university** for about 10 years and writes on his website: "The results of our studies indicate that nanoparticles of titanium dioxide can penetrate the upper layers of the skin and thus also come into contact with living cells. This is not true for zinc oxide, which is why we consider it a safe ingredient for sunscreen."

**Claim in favor of titanium dioxide**

Mr. Peterson, an **industrially funded** professor working in nanoscience in a Swedish **company**, believes that titanium dioxide nanoparticles in particular are expected to penetrate the upper layers of the skin and therefore may have an undesirable effect on our health. The scientist has been working on this topic in his **company** for about 10 years and writes on his website: "Our study results indicate that nanoparticles of titanium dioxide can penetrate the upper layers of the skin and thus also come into contact with living cells. This is not the case for zinc oxide and we therefore consider it a safe ingredient in sunscreen."