

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/373940435>

You won't believe it! Truth judgments for clickbait headlines benefit (but less so) from prior exposure

Article in *Applied Cognitive Psychology* · September 2023

DOI: 10.1002/acp.4134

CITATIONS

0

READS

82

3 authors:



Jérémy Béna

Aix-Marseille Université

20 PUBLICATIONS 85 CITATIONS

SEE PROFILE



Martin Rouard

Université Catholique de Louvain - UCLouvain

1 PUBLICATION 0 CITATIONS

SEE PROFILE



Olivier Corneille

Université Catholique de Louvain - UCLouvain

176 PUBLICATIONS 6,136 CITATIONS

SEE PROFILE

<https://doi.org/10.1002/acp.4134>

**You won't believe it! Truth judgments for clickbait headlines benefit (but less so) from
prior exposure**

Jérémy Béné^{1*}, Martin Rouard¹, & Olivier Corneille¹

¹UCLouvain, Belgium

***Manuscript accepted at Applied Cognitive Psychology. The final version may differ from
this version. doi: 10.1002/acp.4134***

Funding: This work was supported by an FSR Incoming Postdoctoral Fellowship [IPSY FSR22 MOVE] awarded to Jérémy Béné.

Declaration of competing interests: The authors report that there are no competing interests to declare.

Data availability statement: The data that support the findings of the two studies are openly available on the Open Science Framework at

https://osf.io/kecrx/?view_only=710b0f43bd1147ddb47e56ab4e999382.

*Correspondence concerning this article should be addressed to Jérémy Béné, UCLouvain, PSP IPSY, 10 Place du Cardinal Mercier, 1348, Louvain-la-Neuve, Belgium. Email: jeremy.bena@uclouvain.be

CRedit authorship contribution statement

JB: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. MR: Conceptualization, Formal analysis, Investigation, Methodology, Resources, Software, Writing – original draft. OC: Conceptualization, Methodology, Supervision, Writing – review & editing.

Abstract

In two high-powered experiments, we investigated how prior exposure to statements presented in a clickbait format increases the perceived truth of their content. In Experiment 1 ($N = 241$), we hypothesized and found that prior exposure increased the proportion of “true” judgments for both non-clickbait and clickbait content, but with a reduced effect of prior exposure for statements originally presented in a clickbait format. In Experiment 2 ($N = 291$), turning to continuous ratings, we found higher truth ratings for repeated than new clickbait statements, even when repetition evidently originated from prior exposure to clickbait statements. The present findings suggest that exposure to clickbait headlines can increase their content's truth judgments despite their overall lack of credibility, although to a lesser extent than for more regular statements. The present research additionally supports an implausibility account rather than a source memory account of the truth effect with clickbait statements.

Keywords: Clickbait; Headlines; Internet; News; Truth-by-repetition

You won't believe it! Truth judgments for clickbait headlines benefit (but less so) from prior exposure

Online, competition for attention is fierce. When creating or sharing news, several tools can be used to attract attention. One of these tools is clickbait, defined as "content whose main purpose is to attract attention and encourage visitors to click on a link to a particular web page" (Chen et al., 2015, p.1). Because clickbait content can often be misleading (e.g., by exaggerating facts), tools are being developed for its detection (e.g., Chen et al., 2015). The psychological effects of clickbait content remain largely understudied, although some studies have started investigating the effect of clickbait on curiosity (Pengnate, 2019), perceived credibility (e.g., Hurst, 2016), and news understanding and memory (Carcioppolo et al., 2021; Munger et al., 2020; Vultee et al., 2020).

In the current research, we investigated whether prior exposure to clickbait headlines increases the perceived truth of the clickbait content. The truth effect procedure is convenient for examining this question. In a typical truth effect procedure, participants are exposed to a set of unknown factual statements (e.g., "Bramante was an architect of the renaissance") that are then repeated and mixed with new ones in a truth judgment task. The typical result is that statements repeated between the exposure phase and the truth judgment phase receive higher truth judgments than new statements – i.e., a "truth effect" (for a meta-analysis, see Dechêne et al., 2010; for early demonstrations, see Hasher et al., 1977; Bacon, 1979; Schwartz, 1982; for a recent overview, see Unkelbach et al., 2019).

Common explanations of the truth effect involve processing fluency (Reber & Schwarz, 1999). Being repeatedly exposed to a piece of information makes it subsequently easier to process, and this fluency is then used as a cue for truth (e.g., Reber & Scharz, 1999; Unkelbach & Greifeneder, 2013; see Alter & Oppenheimer, 2009, for the influence of fluency on several judgments). Other explanations have been suggested (recognition memory: Bacon,

1979; coherent references in memory: Unkelbach & Rom, 2017); and these explanations are sometimes conceptualized as compatible rather than in competition with fluency accounts (e.g., Unkelbach et al., 2019).

The truth effect is robust across experimental settings (Dechêne et al., 2010), even if some conditions cancel it out (e.g., asking for truth judgments both in the exposure and judgment phases when only a short time interval separates them; see, e.g., Nadarevic & Erdfelder, 2014). Although the truth effect has long been believed to only occur with unknown factual statements (e.g., Dechêne et al., 2010), recent research showed that the effect is, in fact, far more general. For instance, the truth effect was found with true Covid19-related statements (Unkelbach & Speckmann, 2021), opinions (Arkes et al., 1989), rumors (DiFonzo et al., 2016), fake news (Pennycook et al., 2018), conspiracy statements (Béna et al., 2023), statements known to be false by the participants (Fazio et al., 2015; Fazio, 2020), and with highly implausible statements (Lacassagne et al., 2022; see Pillai & Fazio, 2021 for an overview of the effects of prior exposure on false and misleading information). Overall, these results suggest that repeated exposure has a general role in belief formation.

Interestingly, however, prior exposure may not always increase perceived truth. For instance, when statements are introduced as "false" in the exposure phase, repeated statements can be associated with lower truth ratings than new ones provided participants remember the specific "false" source (Begg et al., 1992; Unkelbach & Stahl, 2009). Another example comes from Corneille et al. (2020; see also Béna et al., 2022), who found that repeated statements were more likely to be perceived as "previously used as fake news on social media" than new ones. One possible reason is that when the judgment refers to social media, a context where fabricated and false information may spread widely (e.g., Del Vicario et al., 2016; Vosoughi et al., 2018; Juul & Ugander, 2021), repetition-induced fluency may be a cue for "fake news" judgments rather than for truth. Further investigations, however, are required to better

understand this "fake news by repetition" effect because of interpretational issues with the instructions that were used and because truth effect studies mimicking social media postings found the typical effect (Nadarevic et al., 2020; Smelter & Calvillo, 2020; for a discussion, see Béna et al., 2022).

Based on the above literature, it is unclear whether prior exposure should increase or reverse truth judgments of statements displayed in a clickbait format. On the one hand, clickbait content is typically associated with the Internet, where inaccurate information may spread widely (Marsh & Rajaram, 2019). Clickbait is also often exaggerated, sensationalized information. As a result, displaying clickbait content in the exposure phase could remind participants of misinformation spreading online. If so, prior exposure may decrease truth judgments for the clickbait format (see Béna et al., 2022; Corneille et al., 2020). On the other hand, the truth effect has been found in a wide range of contexts and materials, including misinformation, false, and implausible statements; therefore, the truth effect may generalize to clickbait statements.

In the present research, we predicted that prior exposure would increase truth judgments for statements displayed in a clickbait format, *but to a lesser extent* than for statements displayed in a non-clickbait format. To discuss the rationale for this prediction, a methodological comment is in order. In Experiment 1 below, the clickbait or non-clickbait formats were implemented only in the exposure phase. In the truth judgment task, all statements were presented in a standard format (i.e., as factual neutral trivia statements typically used in truth effect studies). Because clickbait is commonly associated with a lack of credibility, using a clickbait format in the judgment phase would have provided participants with a cue for falsehood; as a result, this judgment cue may have contaminated prior exposure effects. Of note, both clickbait and non-clickbait statements underwent surface changes between the exposure and judgment phases. These surface changes (1) allowed adequate

comparisons between repeated and new statements and (2) ensured that surface changes applied to both clickbait and non-clickbait statements, but (3) they also implied a deviation from the bulk of truth effect studies (in which repetition is typically *verbatim*, i.e., repeating the specific wording). Nevertheless, the truth effect has been found when *thematic* repetition (i.e., repeating the meaning of a statement but not its specific wording) rather than *verbatim* repetition was manipulated (e.g., Begg et al., 1985; Silva et al., 2017). Hence, despite applying surface changes between exposure and judgment to both clickbait and non-clickbait statements, we expected that the effect of repetition should survive. Experiment 2 relied on a more typical truth effect procedure and included only clickbait statements at both the exposure and test phases (see below).

Having clarified this point, we return to the prediction we tested in Experiment 1: the truth effect should be smaller for statements initially presented in a clickbait than in a non-clickbait format. This prediction could be derived from two general models of the truth effect, which we now discuss.

The first account is a *source memory account*, based on Begg et al. (1992; for a general misattribution model, see Jacoby et al., 1989; Kelley & Jacoby, 1990). Begg et al. (1992) found that recollection and familiarity, two retrieval processes, are independent contributors to truth judgments. For instance, participants may judge a given statement as “true” because they remember that it was previously communicated by a reliable source or because the statement feels familiar even if they do not remember its source (i.e., familiarity in the absence of recollection). When a source is introduced as presenting only false statements in the exposure phase, recollection supports a “false” judgment on statements from

this “false” source; in contrast, familiarity¹ in the absence of recollection supports a “true” judgment.

Applying this source memory account to the truth effect with clickbait statements, we reasoned that the clickbait vs. non-clickbait format in the exposure phase might act as source information that participants may or may not remember when judging the truth of repeated and new statements. We assumed that clickbait statements would tend to be perceived as relatively untrustworthy, while non-clickbait statements would tend to be perceived as relatively trustworthy. If participants remember that a statement was originally displayed as a clickbait headline in the exposure phase, this may frequently signal falsehood and support “false” judgments. Yet, participants may not always remember that a statement was originally presented in a clickbait format. In this case, repetition-induced fluency should support “true” judgments. When *non*-clickbait statements are presented in the exposure phase, repetition-induced fluency should support “true” judgments (as for clickbait statements), and recollection should frequently support “true” judgments. As a result, recollection and fluency should generally act in opposition for clickbait statements but generally act in concert for non-clickbait statements. It follows that the truth effect should be smaller for clickbait statements than for non-clickbait statements.

The second account is an *implausibility account*, based on Fazio et al.'s (2019) model. In this model, repeated exposure increases belief in statements (a latent construct) by the same

¹ In the remainder of the manuscript, we use the term “fluency” as it is more widely used in the truth effect literature and close to the notion of familiarity, although familiarity and fluency are two different processes (familiarity can be conceived as resulting from an attribution of fluency to prior exposure).

amount for all statements, regardless of their plausibility. Yet, because highly implausible statements are unlikely to be believed regardless of repeated exposure, a repetition-induced increase in belief (again, defined at the latent level) is unlikely to make statements initially perceived as false appear to be true. Applying this approach to the truth effect with clickbait statements, we reasoned that participants should be more likely to encode a statement as implausible at exposure if it is presented in a clickbait format. In the judgment phase, prior exposure should increase perceived truth, but the statement's initial classification as implausible should make it unlikely that truth judgments on clickbait statements switch from “false” to “true” even if participants do not remember the format used in the exposure phase.

Based on the above rationale, in Experiment 1, we predicted a smaller truth effect with statement contents presented in a clickbait format at exposure. As just discussed, this prediction is consistent with the source memory and implausibility accounts. In Experiment 2, we tested which of these accounts is most suited to explain the truth effect with clickbait statements.

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the two experiments. The preregistrations, experiment programs, data, and analyses are publicly available at

https://osf.io/kecrx/?view_only=710b0f43bd1147ddb47e56ab4e999382.

Experiment 1

In Experiment 1, we tested (1) whether a truth effect is found with statements presented as clickbait headlines in the exposure phase and (2) whether this effect is smaller than that found with statements presented as non-clickbait headlines in the exposure phase.

Methods

Participants and design

We manipulated Repetition within participants: some facts were repeated (displayed both in the exposure phase and in the truth judgment task), and other statements were new (displayed only in the truth judgment task). A nested variable was Format: Repeated factual statements were displayed in a clickbait or non-clickbait format in the exposure phase (within participants). Note that Repetition was not verbatim: statements displayed in a clickbait or non-clickbait format in the exposure phase were displayed as factual neutral statements in the judgment phase. Half of the statements (whether they were repeated [clickbait or non-clickbait] or new) were factually true, and half were factually false.

We collected data online from 241 participants on Prolific (50% females, one not available; $M_{age} = 38.9$; $SD_{age} = 13.55$, three not available). Participants (1) were native English speakers, (2) declared to live in the United States, (3) had an approval rate of at least 95%, (4) had already completed at least 100 previous submissions, and (5) did not take part in previous experiments we conducted on Prolific on the truth effect, nor in the pretests on the materials we used in the present experiment (see below). Participants were paid £ 1.50 (US\$ 1.88) for completing the study. We preregistered that we would exclude participants declaring they did not respond seriously; no participants were affected by this criterion. We ended up with one more participant than we planned because we received the data from one participant who took too much time to complete the study regarding Prolific's standards. As a result, Prolific considered that this "timed-out" participant was not included in the total sample and automatically reopened the data collection for one additional participant. Because taking too much time to complete the study was not one of our preregistered exclusion criteria, we decided to include the data of that timed-out participant, resulting in 241 participants rather than 240.

To determine sample size, we set α to .05. We aimed for a statistical power of at least 90% to detect an effect size of $d = 0.2$ in a one-tailed paired samples t -test (dependent variable: truth effect scores, which is the difference between the proportions of "true" judgments for repeated and new statements; independent variable: Format in the exposure phase, clickbait or non-clickbait). An analysis with G*Power (Version 3.1.9.6, Faul et al., 2007) showed that we needed 216 participants. To avoid a final sample smaller than the targeted sample size, we increased this estimate by 10%, resulting in ~240 participants.

Materials

In the present section, we report the procedure for preparing the study materials, which consisted of two studies: a first study to select the statements and a second study to test their perceived trustworthiness.

Creation and selection of the statements

We first created 100 statements (half factually true; half factually false). Fifty statements were aimed at mimicking clickbait headlines (e.g., "The POPE DOESN'T WANT THIS INFORMATION TO LEAK but the Vatican has the HIGHEST CRIME RATE in the world.") and 50 other statements were aimed at mimicking non-clickbait headlines (e.g., "Aged 76, the Romanian pianist Radu Lupu died in April 2022."). We reasoned that we needed to create new statements because using existing clickbait and non-clickbait headlines can be problematic for two reasons. First, using existing headlines makes it possible that participants have seen at least some of the headlines before the experiment, possibly interfering with one of our critical manipulations, which is repetition. Second and more critically, clickbait and non-clickbait headlines may differ on several dimensions, which makes it difficult to interpret any difference between clickbait and non-clickbait headlines. For instance, clickbait headlines may be more emotional than non-clickbait headlines (e.g.,

Pengnate, 2019). To have better control over the statements, we created or selected factual neutral statements (e.g., "Cats can be allergic to humans") and then adapted the statements to mimic clickbait or non-clickbait headlines. We adapted 26 statements from Corneille et al. (2020) and 24 statements from Unkelbach and Rom, (2017), ending up with 50 statements (half true). We further created 50 factual neutral statements (also half true) and then created their clickbait or non-clickbait headline counterpart to have enough statements.

Because we wanted clickbait statements to mimic clickbait headlines and non-clickbait statements to mimic non-clickbait headlines, we conducted a first online study in which we asked 59 participants (different from the main experiments) to indicate whether each of the 100 statements we created was similar to (1) a clickbait headline, (2) a non-clickbait headline, (3) neither of the two. Participants were provided with a definition of "clickbait" to inform their decision. We kept the 20 statements designed to mimic clickbait headlines most perceived as similar to clickbait headlines (half true) and the 20 statements designed to mimic non-clickbait headlines most perceived as similar to non-clickbait headlines (half true). Each statement in a clickbait or non-clickbait format (used in the exposure phase of Experiment 1) has a factual neutral counterpart, which we used in the truth judgment task of Experiment 1. The statements we selected are available at

https://osf.io/wq9sy?view_only=710b0f43bd1147ddb47e56ab4e999382.

Perception of trustworthiness for clickbait and non-clickbait statements.

Based on previous research (Hurst, 2016; Vultee et al., 2020), we expected clickbait headlines to be perceived as less trustworthy than non-clickbait headlines. To test whether this is also the case for the statements we created, we conducted a second online study ($N = 65$, different from those included in the statements selection phase and from the main experiments).

Participants indicated to what extent they found the 40 statements we selected (see above) trustworthy on a 7-point Likert scale (from "Not trustworthy at all" to "Completely

trustworthy"). We conducted a one-tailed paired samples *t*-test on means of perceived trustworthiness as a function of Format (clickbait vs. non-clickbait). The test was one-tailed because we were interested in statements in a clickbait format being perceived as *less* trustworthy than statements in a non-clickbait format. We found that participants rated clickbait statements as less trustworthy ($M = 2.73$; $SD = 0.76$) than non-clickbait statements ($M = 4.89$; $SD = 0.79$), $t(64) = -15.76$, $p < .001$, $d = 1.95$, $90\%CI_d = [1.60, 2.30]$.

Procedure

We programmed the experiment with lab.js (Henninger et al., 2022), and we used JATOS (Lange et al., 2015) to export the experiment to an HTTPS-protected website.

Upon clicking on the study link, participants were told that we were interested in the role of various variables in the evaluation of statements. After providing their informed consent, the exposure phase task started. Participants rated their interest on 20 statements (10 clickbait, 10 non-clickbait; half true). The statements were presented once in random order. Participants used a 6-point Likert scale (from "Very uninteresting" to "Very interesting") to indicate their interest with no time limit. Upon clicking on their response, the next statement was displayed after an intertrial blank screen of 500 milliseconds.

The instructions were as follows: *"Dear participants, you are participating in a study about the role of various variables in the evaluation of statements. In this first part, you will be presented with several statements. These statements will be displayed individually. Your task will be to indicate the extent to which you think the statements are interesting. To do so you will use a scale ranging from "Very uninteresting" to "Very interesting". For each statement, please select the value that most closely represents what you think."*

Right after the exposure phase², participants entered the truth judgment task. Participants had to indicate whether they thought 40 statements (half seen in the exposure phase as clickbait or non-clickbait headlines; half true) were true or false³. The statements were displayed individually in random order. In the judgment task, all the statements were factual neutral statements similar to the ones typically used in truth effect studies.

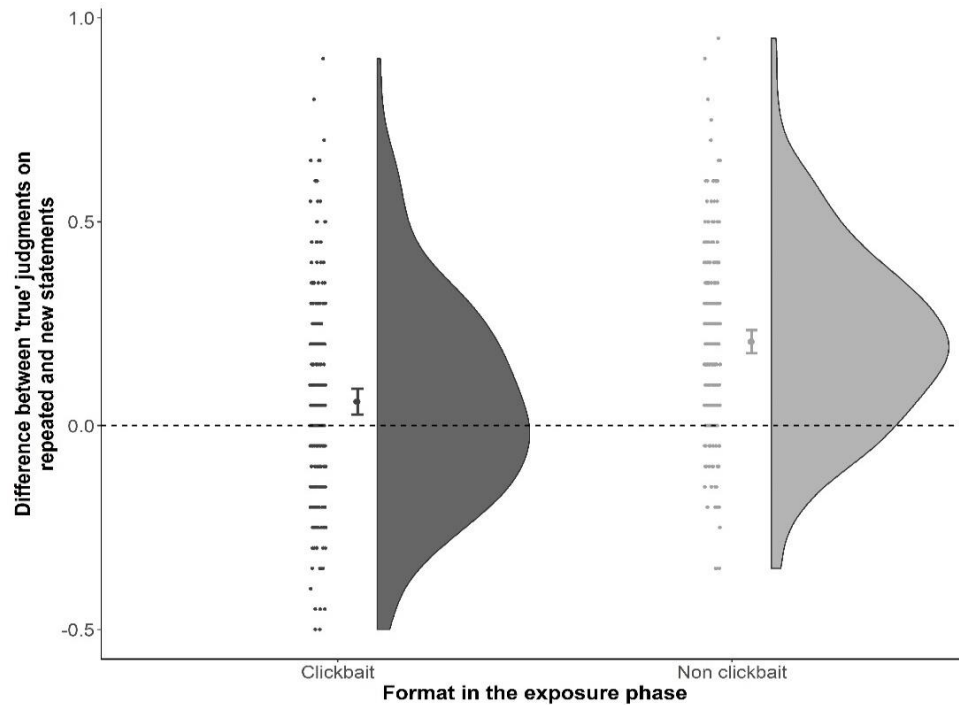
The instructions for the judgment phase were as follows: *"In the following part, you will see several statements. Some are true, others are false. For each statement, please indicate whether you think it is "true" or "false". There will be no time limit. If you think that a statement is true, please click on "True". If you think that a statement is false, please click on "False". When you are ready, please click on 'Continue'."*

After the truth judgment task, participants answered the following question (based on Aust et al., 2013): *"It would be very helpful if you could tell us at this point whether you have taken the requested responses seriously, so that we can use your answers for our scientific*

² Note that we did not include any time interval between the exposure phase and the truth judgment task. When the task in the exposure phase is not truth-related (as the case in the present experiments), the truth effect can be found without any time interval, as documented in several studies (e.g., Corneille et al., 2020; Henderson et al., 2021; Unkelbach & Rom, 2017).

³ Note that we used a dichotomous truth judgment task as it is in line with how we framed the central question we address in Experiment 1 (see Introduction; i.e., does repetition increases “true” judgments for clickbait statements?). Dichotomous truth ratings are routinely used in truth effect research (e.g., Unkelbach & Rom, 2017; Unkelbach & Stahl, 2009), and they are especially relevant when the interest lies in the extent to which repeated exposure increases the proportion of “true” judgments. In some contexts, Likert scales and continuous ratings may also be relevant (e.g., Lacassagne et al., 2022); see the present Experiment 2 for the use of a continuous truth rating scale.

Figure 1. Mean truth effect scores (i.e., proportions of “true” judgments on repeated vs. new statements) as a function of Format (statements featured in Clickbait or Non clickbait headlines in the exposure phase).



Note. The dots are participants' scores (jittered). The error bars are the 95% confidence intervals, with the mean in between. The distributions are the kernel probability density of the data (trimmed to remain within -1; 1, the range of possible values). Dashed horizontal line: no difference between truth judgments on repeated and new statements.

analysis, or whether you were just clicking through to take a look at the survey? (This will not affect your payment)." (response options: *"I have taken the requested responses seriously;" "I have just clicked through, please discard my data"*). Participants were then thanked and debriefed.

Results

Table 1. Means (SD) of truth judgments in each Repetition $\times$ Format condition of Experiment 1 and in each Repetition condition of Experiment 2.

	Experiment 1		Experiment 2
	Clickbait	Non-clickbait	Clickbait
Repeated	0.58 (0.22)	0.73 (0.19)	42.65 (15.34)
New	0.52 (0.16)		40.08 (15.02)
Difference	0.06 (0.25)	0.21 (0.22)	2.57 (13.16)

Note. Truth judgments were provided in a dichotomous truth judgment task in Experiment 1 and on a continuous truth rating scale in Experiment 2. In Experiment 1, statements were presented in a clickbait or non-clickbait format in the exposure phase but not in the judgment phase, where repeated and new statements were presented as neutral trivia statements. In Experiment 2, only clickbait statements were displayed, both in the exposure and test phases. Difference scores reflect the truth effect (truth judgments on repeated content – truth judgments on new statements).

We performed all analyses in R (R Core Team, 2021). For each participant, we computed two truth effect scores⁴: one for factual statements displayed in a clickbait format in the exposure phase, and another for statements displayed in a non-clickbait format in the exposure phase.

⁴ Although non-preregistered and outside the scope of our focal interest, we also calculated the average interest ratings participants provided in the exposure phase. Interest ratings were higher for statements displayed in a non-clickbait format ($M = 3.31$; $SD = 0.79$) than in a clickbait format ($M = 2.84$; $SD = 0.96$), $t(240) = 6.92$, $p < .001$, $d = 0.45$, $95\%CI_d = [0.31; 0.58]$. Although this result may be inconsistent with the idea that clickbait headlines attract attention because they elicit more interest, it is consistent with the possibility that participants are reluctant to provide higher ratings of interest for statements designed to elicit untrustworthiness.

Table 2. Number (and proportion) of participants showing a positive, null, or negative truth effect score and overall mean scores (SD) in each category as a function of Format in Experiment 1 and for clickbait statements in Experiment 2.

	Experiment 1				Experiment 2	
	Clickbait		Non-clickbait		Clickbait	
Truth effect direction	Number (proportion) of participants	Overall effect	Number (proportion) of participants	Overall effect	Number (proportion) of participants	Overall effect
Positive	125 (.52)	.25 (.18)	191 (.79)	.28 (.18)	168 (.58)	11.15 (8.54)
Null	18 (.07)	0	12 (.05)	0	4 (.01)	0
Negative	98 (.41)	-.18 (.11)	38 (.16)	-.12 (.09)	119 (.41)	-9.46 (8.46)

Note. Positive scores are higher truth judgments for repeated vs. new statements; negative scores are higher truth judgments for new vs. repeated statements; null scores reflect no difference. Overall scores are necessarily 0 in the "Null" category. Truth judgments were provided in a dichotomous truth judgment task in Experiment 1 and on a continuous truth rating scale in Experiment 2.

Truth effect scores are the difference between proportions of "true" judgments on repeated statements and proportions of "true" judgments on new statements. Proportions of "true" judgments are typically calculated in truth effect studies (e.g., Unkelbach & Rom, 2017). The truth effect scores are displayed in Figure 1 (see also Table 1 for proportions in each Repetition condition). Positive scores indicate higher truth judgments for repeated than new statements; negative scores indicate higher truth judgments for new than repeated statements; null scores indicate no difference.

First, we conducted a two-tailed one-sample *t*-test ($\mu = 0$) on the truth effect scores computed with statements displayed in a non-clickbait format in the exposure phase. Factual statements that were displayed in a non-clickbait format in the exposure phase were more

often judged as "true" than new statements, ending up in a positive mean truth effect score, $t(240) = 14.25, p < .001, d = 0.92, 95\%CI_d = [0.77; 1.07]$ (see Figure 1).

We repeated the same analysis on truth effect scores calculated with statements displayed in a clickbait format in the exposure phase. Here too, we found higher truth judgments for repeated than new statements, $t(240) = 3.63, p < .001, d = 0.23, 95\%CI_d = [0.11; 0.36]$.

Repetition increased the truth judgments regardless of the statements' format in the exposure phase. As may be apparent above, the effect sizes are different: when statements were displayed in a non-clickbait format, the effect size was larger (truth effect score: $M = .21; SD = .22$) than when they were displayed in a clickbait format (truth effect score: $M = .06; SD = .25$). To test whether this difference was statistically significant, we conducted a preregistered one-sided paired samples t -test on truth effect scores as a function of Format. Supporting our central prediction, truth effect scores were larger in the clickbait than in the non-clickbait condition (see Figure 1), $t(240) = -10.63, p < .001, d = 0.68, 90\%CI_d = [0.57; 0.80]$.

The smaller truth effect scores observed in the clickbait than non-clickbait format condition can originate in several (non-exclusive) patterns, including smaller but still positive effects of Repetition on truth judgments and more participants showing a negative effect of Repetition on truth judgments. As exploratory (and non-preregistered) analyses, we calculated for each participant whether their truth effect score was positive (higher truth judgments for repeated vs. new statements), null (no difference), or negative (higher truth judgments for new vs. repeated statements), separately in the clickbait and non-clickbait Format conditions (see Table 2 for the proportions of participants and the overall score in each category).

Interestingly, approximately 41% of the participants showed a negative truth effect in the clickbait format condition, while this proportion was only about 16% in the non-clickbait

format condition. When we restricted analyses to participants with a positive truth effect score both in the clickbait and non-clickbait format conditions ($n = 114$), we found larger truth effect scores in the non-clickbait ($M = .35$; $SD = .18$) than clickbait Format condition ($M = .26$; $SD = .18$), $t(113) = -5.34$, $p < .001$, $d = 0.50$, 95% $CI_d = [0.30; 0.69]$. These additional analyses suggest two patterns for why the truth effect was smaller in the clickbait than in the non-clickbait Format condition: More participants having negative truth effect scores and smaller positive truth effect scores.

Discussion

In Experiment 1, we tested whether prior exposure to statements presented as clickbait headlines would increase the perceived truth of their content (when compared to new statements – i.e., the truth effect) and whether this truth effect would be smaller than for statements displayed as non-clickbait headlines in the exposure phase. The findings are consistent with the pattern we predicted: we found a truth effect both for statement contents presented in a clickbait and non-clickbait format in the exposure phase, but a smaller one for the former. As explained in the introduction, these findings are consistent with both the source memory and implausibility accounts. However, the two accounts may not always align. In particular, they make incompatible predictions in specific experimental settings. To gain insight into the underlying processes of the truth effect with clickbait statements, we conducted Experiment 2.

Experiment 2

Experiment 2 aimed at contrasting the source memory and implausibility accounts of the truth effect with clickbait statements regarding one specific prediction they make. For the source memory account, source recollection of the (clickbait or non-clickbait) format is critical, as participants should not misattribute repetition-induced fluency to truth when they remember

that the statement was presented earlier in an (untrustworthy) clickbait format. In contrast, the implausibility account leaves no role for statement format memory: when a statement is presented in a clickbait format, participants should be likely to encode it as implausible. In the judgment phase, prior exposure should increase perceived truth, but the statement's initial classification as implausible should make it unlikely that truth judgments on clickbait statements switch from “false” to “true,” irrespective of source memory. As it appears, source memory is key for the source memory account but not for the implausibility account.

We reasoned that testing the truth effect with statements presented in a clickbait format both in the exposure and test phases, for both repeated and new statements, in conjunction with the use of a sensitive truth rating scale (substituting continuous measures to the dichotomous judgments used in Experiment 1), would allow testing the source memory and implausibility accounts against each other.

For the source memory account, if all statements displayed in the exposure and judgment phases are clickbait statements, participants should know that all the statements they recognize in the judgment phase come from an untrustworthy source, especially when the judgment task comes right after the exposure phase, as this was the case in this second experiment. Knowing that all the repeated statements are untrustworthy should make participants discount repetition-induced fluency as a cue for truth. Hence, based on the source memory account, we expect to find either no truth effect or a reversed effect (i.e., higher truth ratings for new than repeated statements) in the present experiment. Therefore, finding a truth effect despite experimental conditions conducive to high levels of remembrance would challenge the source memory account. In contrast, source memory is irrelevant for the implausibility account: repetition-induced fluency should increase the perceived truth for any statement. Such an increase may not involve a switch from “false” to “true” judgments; it may translate to a small increase that could be observed on a sensitive truth judgment scale (see

Lacassagne et al., 2022). As a result, finding a reversed truth effect or no effect while using a sensitive truth judgment scale would challenge the implausibility account.

Methods

Participants and design

We manipulated Repetition within participants: some statements were repeated (displayed both in the exposure phase and in the truth judgment task), and other statements were new (displayed only in the truth judgment task). Importantly, and in contrast to Experiment 1, all statements were displayed as clickbait statements in both the exposure and judgment phases. Half of the statements (whether repeated or new) were factually true, and half were factually false.

We collected data online from 292 participants on Prolific (50% females; $M_{age} = 41.35$; $SD_{age} = 14.65$, eight not available). Participants (1) were native English speakers, (2) declared to live in the United States, (3) had an approval rate of at least 95%, (4) had already completed at least 100 previous submissions, and (5) did not take part in previous experiments we conducted on Prolific on the truth effect, nor in the pretests and Experiment 1 we reported above. Participants were paid £ 0.75 (US\$ 0.94) for completing the study. We excluded one participant who declared they did not respond seriously, ending up with a sample size of 291 participants.

To determine this sample size, we set α to .05, and we aimed for a statistical power of at least 90% to detect an effect size of Cohen's $d = 0.20$ in a (two-tailed) paired samples t -test (DV: mean truth ratings; IV: Repetition). We chose Cohen's $d = 0.20$ as this is roughly the size of the truth effect we found with clickbait statements in Experiment 1 ($d = 0.23$). An analysis with G*power (version 3.1.9.6; Faul et al., 2007) showed that we needed 265

participants. To avoid having a final sample smaller than the targeted sample size, we increased this estimate by 10%, resulting in ~292 participants.

Materials and procedure

Experiment 2 departs from Experiment 1 in four regards. First, we only used the 20 statements (half factually true) that were used as “clickbait” headlines in Experiment 1. Second, repeated and new statements in the truth judgment task were all displayed in a clickbait format. Third, 10 statements were displayed in the exposure phase (randomly selected from the set of 20 statements), and 20 statements were displayed in the judgment phase. Fourth, we replaced the dichotomous responses (“True”; “False”) used in Experiment 1 with a continuous rating scale (between 0 “Definitely false” and 100 “Definitely true”) to detect possibly small differences in rated truth between repeated and new statements that do not involve a qualitative decision on the “false” vs. “true” continuum.

Results

We calculated the mean truth ratings for each participant separately for repeated and new statements (see Table 1). Higher ratings indicate higher perceived truth.

As preregistered, we conducted a frequentist and a default Bayesian two-tailed paired samples *t*-test on the truth ratings as a function of Repetition. Critical for our main test, truth ratings were higher for repeated than new statements, $t(290) = 3.33$, $p = .001$, $d = 0.20$, $95\%CI_d = [0.08; 0.31]$. The Bayesian analysis yielded evidence for this effect, $BF_{10} = 14.42$.

As in Experiment 1, we also calculated the qualitative direction of the truth effect scores (i.e., truth ratings on repeated vs. new statements) for each participant (see Table 2 for the proportions of participants and the overall scores in each category). These analyses are exploratory and were not preregistered. Approximately 41% of the participants showed a negative truth effect in the clickbait format condition, which is identical to the proportion we

found in Experiment 1 despite (1) changes in the measure, (2) presenting only clickbait statements in the exposure phase (rather than clickbait and non-clickbait statements as in Experiment 1), and (3) presenting statements in a clickbait format in the truth rating task (instead of presenting statements in a clickbait format only in the exposure phase).

Discussion

In Experiment 2, we sought to gain insight into the underlying processes involved in the truth effect with clickbait statements by opposing two accounts that converged in the prediction we tested in Experiment 1, but that diverged in predictions made for Experiment 2. For the source memory account, remembering that a statement was presented in the exposure phase should make participants discount repetition-induced fluency (as all statements were presented as clickbait statements in the exposure phase). In contrast, for the implausibility account, there is no role for source memory: repetition should increase perceived truth even for clickbait statements which, as we argued, participants may see as relatively implausible.

To increase the likelihood that participants remember that the repeated statements were presented in the exposure phase and that these statements are clickbait (untrustworthy) statements, we (1) used only clickbait statements in the exposure phase, (2) used a verbatim repetition, meaning that the statements were also displayed as clickbait statements in the truth judgment task, (3) used few statements (20 statements; half repeated), (4) administered the truth rating task right after the exposure phase. Verbatim repetition, together with short lists of statements and no time interval, are conducive to detailed and conscious memories (e.g., Brainerd et al., 2003; Gardiner & Java, 1991; Yonelinas, 1994, 2002). In addition, participants used a continuous truth rating scale so that we could detect even a small truth effect that does not imply a qualitative switch from “false” to “true” responses (for a similar rationale, see, e.g., Lacassagne et al., 2022). With such experimental settings, we expected to find either no

truth effect or even a reversed effect based on the source memory account but a truth effect based on the implausibility account.

The results are more compatible with the implausibility account than the source memory account: we found higher truth ratings for repeated than new clickbait statements. While these results are in line with the prediction we derived from the implausibility account, they are in contradiction with the source memory account. For this reason, the present findings suggest that the (small) truth effect observed with clickbait statements more likely results from repetition-induced fluency attributed to truth even when participants perceive clickbait statements as relatively implausible, rather than from discounting fluency as a cue for truth because participants remember that only clickbait, untrustworthy, statements were displayed in the exposure phase.

General Discussion

Clickbait is a common online journalism strategy aimed at attracting attention and clicks; it does not aim to be accurate and is associated with low perceived trustworthiness. Despite its overall lack of credibility, does the presentation of statements in a clickbait format increase subsequent truth judgments about their content? Based on two models of the truth effect (i.e., higher truth judgments for repeated than new statements, see, e.g., Dechêne et al., 2010), we predicted that the truth effect might extend to clickbait content, although to a lesser extent than for non-clickbait content.

To test this hypothesis, we first created statements that mimicked clickbait and non-clickbait headlines. We selected the statements perceived as most similar to clickbait and non-clickbait headlines. Based on previous research (Hurst, 2016; Vultee et al., 2020), we assumed that statements would be perceived as less trustworthy when displayed in a clickbait

than in a non-clickbait format. This is what we found when pretesting the materials (see section "*Perception of trustworthiness for clickbait and non-clickbait statements*").

In Experiment 1, repeated statements were displayed in a clickbait or non-clickbait format in the exposure phase and mixed with new ones in a dichotomous truth judgment task. Consistent with predictions, we found a truth effect with both clickbait and non-clickbait statements, but a smaller one for the clickbait ones. Of interest, complementary analyses indicated that this difference originates in a higher prevalence of *negative* truth effects and in smaller positive truth effects for clickbait statements.

In Experiment 2, we used (1) only clickbait statements in the exposure and judgment phases, (2) an experimental setting typically conducive to high levels of remembering, and (3) a continuous truth rating scale. This was done to examine whether the effect of repetition on truth judgments for clickbait statements would be more consistent with (1) repetition-induced fluency attributed to truth even when participants perceive clickbait statements as relatively implausible (implausibility account; predicting a positive effect of repetition) or (2) by discounting fluency as a cue for truth because participants remember that only clickbait, untrustworthy, statements were displayed in the exposure phase (source memory account; predicting a null or even a reversed effect of repetition). Since we found a truth effect, the findings better support the implausibility account than the source memory account.

Contrary to the bulk of truth effect studies, Experiment 1 applied surface changes to both clickbait and non-clickbait statements between the exposure phase and the truth judgment task. The truth effect is well replicated when the exact same statements are used in the exposure phase and the truth judgment task (Dechêne et al., 2010). Research is more scarce, however, regarding the truth effect when *thematic* repetition rather than *verbatim* repetition is involved. The present results contribute to further documenting the truth effect despite alterations in statements' exact form, in line with previous research (Begg et al., 1985;

Silva et al., 2017). Of note, we found the truth effect with clickbait statements, whether repetition was *verbatim* (Experiment 2) or based on changes in the statements (Experiment 1).

The truth effect we found for clickbait statements in the two experiments dovetails nicely with recent research showing the truth effect with false and misleading information (Pillai & Fazio, 2021), such as fake news (Pennycook et al., 2018), implausible statements (Fazio et al., 2019; Lacassagne et al., 2022), statements known to be false (Fazio et al., 2015; Fazio, 2020), and conspiracy theories (Béna et al., 2023). However, the present study designs differ from these past studies, for the content of the statements we used is not epistemically dubious. Instead, credibility information was conveyed by the framing (i.e., clickbait or non-clickbait format). In this sense, the present studies resemble research where credibility information is provided through sources in the exposure phase (e.g., Begg et al., 1992; Henkel & Mattson, 2011; Unkelbach & Stahl, 2009).

As just discussed, the findings of Experiment 2 better support the implausibility account (based on Fazio et al., 2019) than the source memory account (based on, e.g., Begg et al., 1992). However, this conclusion does not imply that Fazio et al.'s model *in general* is better supported than a memory misattribution approach. This is because we did not proceed to test these models; rather, we tested predictions we derived from accounts inspired by these models when applied to a clickbait scenario. Finding the truth effect in Experiment 2 contradicts the prediction we derived from the source memory account (i.e., no truth effect or a reversed truth effect). Yet, this does not mean that individuals do not sometimes rely on source recollection (of credibility information) and sometimes rely on familiarity/fluency when recollection fails to judge the truth of information. Instead, it means that such a processing account of how the truth effect occurs with clickbait statements is unsatisfactory. If one's aim is to oppose general processing models of the truth effect, we believe this should

be done by testing central assumptions and predictions of these models, which was not the goal of the current endeavor.

Experiment 2 aimed at opposing two accounts of the truth effect with clickbait statements that critically diverge regarding the role of memory for statements' format in the exposure phase. Yet, we did not estimate memory processes (which, under specific conditions, is possible to achieve in truth judgment tasks; see, e.g., Begg et al., 1992; Nadarevic & Erdfelder, 2019; Niedziałkowska & Nieznański, 2021; Unkelbach & Stahl, 2009). As noted in the Discussion of Experiment 2, we used experimental settings typically conducive to high levels of conscious memory (recollection) because we reasoned this would allow a comparison of the source memory account with the implausibility account. However, the test we conducted is based on the assumption of high levels of recollection, which is not documented in the present experiments. We believe this assumption is reasonable considering the current settings, but it could be directly tested in future research in the context of truth judgments of clickbait statements. In particular, the Process Dissociation Procedure (e.g., Begg et al., 1992; Yonelinas & Jacoby, 2012) and Multinomial Processing Tree models (Unkelbach & Stahl, 2009) could be used to obtain estimates of familiarity and source recollection in truth judgments. Conceptually, Experiment 1 presents some similarities with this approach. The non-clickbait format condition resembles an inclusion condition. This is because we hypothesized that recollection should tend to support "true" judgments (to the extent that non-clickbait statements are perceived as relatively trustworthy, which we found in a pretest; see section *Perception of trustworthiness for clickbait and non-clickbait statements* of Experiment 1), and that familiarity should also support "true" judgments. The clickbait format condition resembles an exclusion condition, as we hypothesized that recollection should tend to support "false" judgments (to the extent that clickbait statements are perceived as relatively untrustworthy, which we found in a pretest; see section *Perception of*

trustworthiness for clickbait and non-clickbait statements), and that familiarity should support “true” judgments (recollection and familiarity would generally support opposite responses as a result).

Despite this similarity between Experiment 1 and the Process Dissociation Procedure logic, we deemed Experiment 1 unsuitable to deliver recollection and familiarity estimates. First, this approach would be based on undocumented assumptions in the present clickbait context (e.g., process invariance assumptions, for which the contribution of recollection and familiarity are the same in the inclusion and exclusion conditions). Second, obtaining parameter estimates requires that recollection supports specific responses in the inclusion and exclusion conditions. However, although recollection should *tend* to support “true” judgments for non-clickbait statements and “false” judgments for clickbait statements, this might not always be the case (e.g., some participants may doubt that some non-clickbait statements are true because non-clickbait headlines are sometimes inaccurate). Even if the format (clickbait or non-clickbait) acts as credibility information, it is not as straightforward as source credibility manipulations implemented in previous truth effect studies that used the Process Dissociation Procedure (e.g., introducing statements from Source A as true and statements from Source B as false).

Besides adapting the Process Dissociation Procedure and Multinomial Processing Tree models, several approaches could be used to gain further insights into memory processes underlying the truth effect with clickbait statements. We distinguish between an experimental approach and a correlational approach. In an experimental approach, one could manipulate factors known to decrease source recollection without substantially affecting familiarity in experiments similar to Experiment 1. This would allow testing whether such a manipulation increases the truth effect with clickbait statements. For instance, one such factor could consist in dividing attention at exposure (e.g., Gardiner & Parkin, 1990; Jacoby et al., 1993; see

Yonelinas, 2002), which has been found to increase the truth effect with statements presented by a “false” source in the exposure phase (Begg et al., 1992).

In a more correlational approach, direct memory tasks, such as recognition tasks, could be administered in conjunction with truth judgment tasks in truth effect procedures (as sometimes done in truth effect research). However, simple recognition tasks alone provide no estimates of memory processes, as familiarity and source recollection both typically support “old” responses. One possibility is to complement recognition tasks with a Remember/Know paradigm (e.g., Gardiner, 1988; Tulving, 1985; but see Haaf et al., 2021; in the context of the truth effect, see Law, 1998) to have access to retrieval phenomenology. Such an approach, which could be combined with the experimental approach mentioned above, would allow investigating the relationship between “true” judgments on clickbait versus non-clickbait statements and the phenomenology at retrieval as estimated in a separate, direct memory task.

Turning to a more applied perspective, clickbait is widely used to attract attention (Bazaco et al., 2019), but its psychological effects have been scarcely (and not systematically) investigated. Previous research has examined, for instance, the effect of clickbait on curiosity and arousal (Pengnate, 2019), perceived credibility (e.g., Hurst, 2016), and news understanding and memory (Carcioppolo et al., 2021; Munger et al., 2020; Vultee et al., 2020). To our knowledge, however, no prior research has investigated the effect of repeated exposure to clickbait, despite the prevalence of copying found, for instance, in online news production (e.g., Cagé et al., 2020), a production that is associated with the use of clickbait strategies (Bazaco et al., 2019). As a result, repeated exposure to clickbait content is likely to be common. The present results suggest that despite their overall lack of trustworthiness, truth judgments of clickbait content benefit from repetition. The findings of Experiment 2 suggest that this could be the case even when it is evident that the content comes from clickbait headlines. Hence, even if clickbait headlines are not designed to be accurate but to attract

attention and clicks, their repeated exposure may increase the perceived truth of their content, possibly outside the laboratory.

Despite their possible relevance outside the laboratory, the current findings should not be overgeneralized. We relied on experimental settings that are unlikely to capture some aspects of how clickbait headlines can be encountered. For instance, in the present experiments, there was no time interval between the exposure and test phases, while such immediacy is probably lacking in daily life events (e.g., when people see a clickbait headline and only later on, possibly days after, reencounter similar content). In addition, the experimental settings we used are deprived of rich sources of information that may be integrated to judge the truth of information (e.g., source information; images presented with statements; reactions and comments available on social media). Another possible difference between the present experiments and everyday life contexts is that we created clickbait statements by formatting neutral trivia statements. As a result, the current investigation may not generalize to contents presented in clickbait headlines in online news. Future research may benefit from integrating more ecologically relevant factors into experiments to better estimate the generalizability of the present findings to everyday situations.

Conclusion

Clickbait headlines do not aim for accuracy, and people may know that their content is not trustworthy. Yet, being repeatedly exposed to clickbait, a strategy to attract attention on the Internet, can increase the perceived truth of the information conveyed in clickbait headlines. We reported two experiments (1) showing that a single exposure to clickbait statements increased the likelihood of judging their content as "true" when subsequently encountered, although to a lesser extent than for non-clickbait statements, and (2) suggesting that the truth effect with clickbait statements could occur even when repetition evidently originated from

prior exposure to clickbait statements. The present findings suggest that repeated exposure to clickbait content may be consequential for belief formation, calling for more systematic research on the effects of clickbait exposure.

References

- Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the tribes of fluency to form a metacognitive nation. *Personality and Social Psychology Review*, 13(3), 219-235.
<https://doi.org/10.1177/1088868309341564>
- Arkes, H. R., Hackett, C., & Boehm, L. (1989). The generality of the relation between familiarity and judged validity. *Journal of Behavioral Decision Making*, 2, 81-94.
- Aust, F., Diedenhofen, B., Ullrich, S., & Musch, J. (2012). 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>
- Bacon, F. T. (1979). Credibility of repeated statements: Memory for trivia. *Journal of Experimental Psychology: Human Learning and Memory*, 5(3), 241–252.
<https://doi.org/10.1037/0278-7393.5.3.241>
- Bazaco, A., Redondo, M., Sánchez-García, P. (2019): "Clickbait as a strategy of viral journalism: conceptualization and methods". *Revista Latina de Comunicación Social*, 74, 94-115. <https://doi.org/10.4185/RLCS-2018-1323en>
- Begg, I. M., Anas, A., & Farinacci, S. (1992). Dissociation of processes in belief: Source recollection, statement familiarity, and the illusion of truth. *Journal of Experimental Psychology: General*, 121(4), 446-458. <https://doi.org/10.1037/0096-3445.121.4.446>
- Begg, I., Armour, V., & Kerr, T. (1985). On believing what we remember. *Canadian journal of behavioural science*, 17(3), 199-214.
- Béna, J., Rihet, M., Carreras, O., & Terrier, P. (2023). Repetition could increase the perceived truth of conspiracy theories. *Psychonomic Bulletin & Review*.
<https://doi.org/10.3758/s13423-023-02276-4>

- Béna, J., Corneille, O., Mierop, A., & Unkelbach, C. (2022). Robustness tests provide further support for an ecological account of the truth and fake news by repetition effects. *International Review of Social Psychology*, 35(1), 19. <https://doi.org/10.5334/irsp.683>
- Brainerd, C. J., Reyna, V. F., Wright, R., & Mojardin, A. H. (2003). Recollection rejection: False-memory editing in children and adults. *Psychological Review*, 110(4), 762–784. <https://doi.org/10.1037/0033-295x.110.4.762>
- Carcioppolo, N., Lun, D., & McFarlane, S. (2021). Exaggerated and Questioning Clickbait Headlines and Their Influence on Media Learning. *Journal of media psychology*, 34(1). <https://doi.org/10.1027/1864-1105/a000298>
- Cagé, J., Hervé, N., & Viaud, M.-L. (2020). The production of information in an online world. *The review of economics studies*, 87(5), 2126-2164. <https://doi.org/10.1093/restud/rdz061>
- Chen, Y., Conroy, N. J., & Rubin, V. L. (2015). Misleading online content: Recognizing clickbait as "false news". *WMDD '15: Proceedings of the 2015 ACM on Workshop on Multimodal Deception Detection*, 15-19. <https://doi.org/10.1145/2823465.2823467>
- Corneille, O., Mierop, A., & Unkelbach, C. (2020). Repetition increases both the perceived truth and fakeness of information: An ecological account. *Cognition*, 205, 104470. <https://doi.org/10.1016/j.cognition.2020.104470>
- Dechêne, A., Stahl, C., Hansen, J., & Wänke, M. (2010). The Truth About the Truth: A Meta-Analytic Review of the Truth Effect. *Personality and Social Psychology Review*, 14(2), 238–257.
- Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., ... Quattrociocchi, W. (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences*, 113(3), 554–559. <https://doi.org/10.1073/pnas.1517441113>

- DiFonzo, N., Beckstead, J., Stupak, N., & Walders, K. (2016). Validity judgments of rumors heard multiple times: The shape of the truth effect. *Social Influence*, 11, 118.
- Fazio, L. K. (2020). Repetition Increases Perceived Truth Even for Known Falsehoods. *Collabra : Psychology*, 6(38). <https://doi.org/10.1525/collabra.347>
- Fazio, L. K., Brashier, N. M., Payne, B. K., & Marsh, E. J. (2015). Knowledge does not protect against illusory truth. *Journal of Experimental Psychology: General*, 144(5), 9931002. <https://doi.org/10.1037/xge0000098>
- Fazio, L. K., Rand, D. G., & Pennycook, G. (2019). Repetition increases perceived truth equally for plausible and implausible statements. *Psychonomic Bulletin & Review*, 26(5), 17051710. <https://doi.org/10.3758/s13423-019-01651-4>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/bf03193146>
- Gardiner, J. M. (1988). Functional aspects of recollective experience. *Memory and Cognition*, 16(4), 309–313. <https://doi.org/10.3758/bf03197041>
- Gardiner, J. M., & Parkin, A. J. (1990). Attention and recollective experience in recognition memory. *Memory & Cognition*, 18(6), 579–583. <https://doi.org/10.3758/bf03197100>
- Gardiner, J. M., & Java, R. I. (1991). Forgetting in recognition memory with and without recollective experience. *Memory & Cognition*, 19(6), 617–623. <https://doi.org/10.3758/BF03197157>
- Haaf, J. M., Rhodes, S., Naveh-Benjamin, M., Sun, T., Snyder, H. K., & Rouder, J. N. (2021). Revisiting the remember–know task: Replications of Gardiner and Java (1990). *Memory & Cognition*, 49(1), 46–66. <https://doi.org/10.3758/s13421-020-01073-x>

- Hasher, L., Goldstein, D., & Toppino, T. (1977). Frequency and the conference of referential validity. *Journal of Verbal Learning and Verbal Behavior*, 16(1), 107–112.
- Henderson, E., Simons, D. J., & Barr, D. J. (2021). The Trajectory of Truth: A Longitudinal Study of the Illusory Truth Effect. *Journal of Cognition*, 4(1), 29.
<https://doi.org/10.5334/joc.161>
- Henkel, L. A., & Mattson, M. E. (2011). Reading is believing: the truth effect and source credibility. *Consciousness and cognition*, 20, 1705-1721.
- Henninger, F., Shevchenko, Y., Mertens, U. K., Kieslich, P. J., & Hilbig, B. E. (2022). Lab.js: A free, open, online study builder. *Behavior Research Methods*, 54, 556–573. <https://doi.org/10.3758/s13428-019-01283-5>
- Hurst, N. (2016). *To clickbait or not to clickbait? An examination of clickbait headline effects on source credibility* [Master thesis, University of Missouri-Columbia].
<https://mospace.umsystem.edu/xmlui/bitstream/handle/10355/56080/research.pdf?isAllowed=y&sequence=2>
- Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, 30(5), 513–541.
[https://doi.org/10.1016/0749-596X\(91\)90025-F](https://doi.org/10.1016/0749-596X(91)90025-F)
- Jacoby, L. L., Kelley, C. M., & Dywan, J. (1989). Memory attributions. In H. L. Roediger & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 391–422). New York: Psychology Press.
- Jacoby, L. L., Toth, J. P., & Yonelinas, A. P. (1993). Separating conscious and unconscious influences of memory: Measuring recollection. *Journal of Experimental Psychology. General*, 122(2), 139–154. <https://doi.org/10.1037/0096-3445.122.2.139>

- Juul, J. L., & Ugander, J. (2021). Comparing information diffusion mechanisms by matching on cascade size. *Proceedings of the National Academy of Sciences*, 118(46), e2100786118, <https://doi.org/10.1073/pnas.2100786118>
- Kelley, C. M., & Jacoby, L. L. (1990). The construction of subjective experience: Memory attributions. *Mind & Language*, 5(1), 49-68. <https://doi.org/10.1111/j.1468-0017.1990.tb00152.x>
- Lacassagne, D., Béna, J., & Corneille, O. (2022). Is Earth a perfect square? Repetition increases the perceived truth of highly implausible statements. *Cognition*, 105052. <https://doi.org/10.1016/j.cognition.2022.105052>
- Lange, K., Kühn, S., & Filevich, E. (2015). Correction: "Just Another Tool for Online Studies" (JATOS): An Easy Solution for Setup and Management of Web Servers Supporting Online Studies. *PLOS ONE*, 10(7), e0134073. <https://doi.org/10.1371/journal.pone.0134073>
- Law, S. (1998). Do we believe what we remember or, do we remember what we believe? *Advances in Consumer Research*, 25, 221-225.
- Munger, K., Luca, M., Nagler, J., & Tucker, J (2020). The (null) effects of clickbait headlines on polarization, trust, and learning. *Public opinion quarterly*, 84(1), 49-73. <https://doi.org/10.1093/poq/nfaa008>
- Marsh, E. J., & Rajaram, S. (2019). The digital expansion of the mind: Implications of internet usage for memory and cognition. *Journal of Applied Research in Memory and Cognition*, 8(1), 1–14. <https://doi.org/10.1016/j.jarmac.2018.11.001>
- Nadarevic, L., & Erdfelder, E. (2014). Initial judgment task and delay of the final validity-rating task moderate the truth effect. *Consciousness and Cognition*, 23, 74–84. <https://doi.org/10.1016/j.concog.2013.12.002>

- Nadarevic, L., & Erdfelder, E. (2019). More evidence against the Spinozan model: Cognitive load diminishes memory for “true” feedback. *Memory & Cognition*, 47(7), 1386–1400. <https://doi.org/10.3758/s13421-019-00940-6>
- Nadarevic, L., Reber, R., Helmecke, A. J., & Köse, D. (2020). Perceived truth of statements and simulated social media postings: an experimental investigation of source credibility, repeated exposure, and presentation format. *Cognitive Research: Principles and Implications*, 5(1). <https://doi.org/10.1186/s41235-020-00251-4>
- Niedziałkowska, D., & Nieznański, M. (2021). Recollection of “true” feedback is better than “false” feedback independently of a priori beliefs: an investigation from the perspective of dual-recollection theory. *Memory*, 29(9), 1186–1196. <https://doi.org/10.1080/09658211.2021.1973037>
- Pengnate, S. (2019). Shocking secret you won't believe! Emotional arousal in clickbait headlines. *Online information review*, 43(7), 1136-1150. <https://doi.org/10.1108/OIR-05-2018-0172>
- Pennycook, G., Cannon, T. D., & Rand, D. G. (2018). Prior exposure increases perceived accuracy of fake news. *Journal of Experimental Psychology: General*, 147(12), 1865–1880. <https://10.1037/xge0000465>
- Pillai, R. M., & Fazio, L. K. (2021). The effects of repeating false and misleading information on belief. *WIREs Cognitive Science*, 12(6). <https://doi.org/10.1002/wcs.1573>
- Reber, R., & Schwarz, N. (1999). Effects of Perceptual Fluency on Judgments of Truth. *Consciousness and Cognition*, 8(3), 338–342. <https://doi.org/10.1006/ccog.1999.0386>
- R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: <https://www.R-project.org/>

- Schwartz, M. (1982). Repetition and rated truth value of statements. *The American Journal of Psychology*, 95(3), 393-407.
- Silva, R. R., Garcia-Marques, G., & Reber, R. (2017). The informative value of type of repetition: Perceptual and conceptual fluency influences on judgments of truth. *Consciousness and cognition*, 51, 53-67.
- Smelter, T. J., & Calvillo, D. P. (2020). Pictures and repeated exposure increase perceived accuracy of news headlines. *Applied Cognitive Psychology*, 34(5), 1061–1071.
<https://doi.org/10.1002/acp.3684>
- Tulving, E. (1985). Memory and consciousness. *Canadian Psychology*, 26(1), 1–12.
<https://doi.org/10.1037/h0080017>
- Unkelbach, C., Koch, A., Silva, R. R., & Garcia-Marques, T. (2019). Truth by Repetition: Explanations and Implications. *Current Directions in Psychological Science*, 28(3), 247-253. <https://doi.org/10.1177/0963721419827854>
- Unkelbach, C., & Greifeneder, R. (2013). A general model of fluency effects in judgment and decision making. In C. Unkelbach & R. Greifeneder (Eds.), *The experience of thinking: How the fluency of mental processes influences cognition and behaviour* (pp. 11–32). New York: Psychology Press.
- Unkelbach, C., & Rom, S. C. (2017). A referential theory of the repetition-induced truth effect. *Cognition*, 160, 110–126. <https://doi.org/10.1016/j.cognition.2016.12.016>
- Unkelbach, C., & Speckmann, F. (2021). Mere repetition increases belief in factually true COVID-19-related information. *Journal of Applied Research in Memory and Cognition*, 10(2), 241–247. <https://doi.org/10.1016/j.jarmac.2021.02.001>

- Unkelbach, C., & Stahl, C. (2009). A multinomial modeling approach to dissociate different components of the truth effect. *Consciousness and Cognition*, 18(1), 22-38.
<https://doi.org/10.1016/j.concog.2008.09.006>
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 11461151. <https://doi.org/10.1126/science.aap9559>
- Vultee, F., Burgess, G. S., Frazier, D., & Mesmer, K. (2020). Here's What to Know About Clickbait: Effects of Image, Headline and Editing on Audience Attitudes. *Journalism practice*, 16(1), 1-18. <https://doi.org/10.1080/17512786.2020.1793379>
- Yonelinas, A. P. (1994). Receiver-operating characteristics in recognition memory: Evidence for a dual-process model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(6), 1341–1354. <https://doi.org/10.1037/0278-7393.20.6.1341>
- Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of memory and language*, 46(3), 441-517.
<https://doi.org/10.1006/jmla.2002.2864>
- Yonelinas, A. P., & Jacoby, L. L. (2012). The process-dissociation approach two decades later: Convergence, boundary conditions, and new directions. *Memory & Cognition*, 40(5), 663–680. <https://doi.org/10.3758/s13421-012-0205-5>