



Monetary incentives do not reduce the repetition-induced truth effect

Felix Speckmann¹ · Christian Unkelbach¹

Accepted: 23 November 2021 / Published online: 16 December 2021
© The Author(s) 2021

Abstract

People rate and judge repeated information more true than novel information. This truth-by-repetition effect is of relevance for explaining belief in fake news, conspiracy theories, or misinformation effects. To ascertain whether increased motivation could reduce this effect, we tested the influence of monetary incentives on participants' truth judgments. We used a standard truth paradigm, consisting of a presentation and judgment phase with factually true and false information, and incentivized every truth judgment. Monetary incentives may influence truth judgments in two ways. First, participants may rely more on relevant knowledge, leading to better discrimination between true and false statements. Second, participants may rely less on repetition, leading to a lower bias to respond "true." We tested these predictions in a preregistered and high-powered experiment. However, incentives did not influence the percentage of "true" judgments or correct responses in general, despite participants' longer response times in the incentivized conditions and evidence for knowledge about the statements. Our findings show that even monetary consequences do not protect against the truth-by-repetition effect, further substantiating its robustness and relevance and highlighting its potential hazardous effects when used in purposeful misinformation.

Keywords Truth effect · Repetition · Fake news · Cognitive illusions · Misinformation · Incentivized responding · Conspiracy theories

People see, read, and hear many different facts and statements each day (e.g., news, social media, conversations), which they can believe or doubt. Apparently, people use repetition as a cue to make this judgment; thus, believing repeated statements more compared with nonrepeated statements, a phenomenon known as the illusory truth effect, a truth-by-repetition effect, or simply a truth effect (Brashier & Marsh, 2020; Unkelbach et al., 2019).

In the seminal work by Hasher et al. (1977), the authors presented participants with 60 statements in three different sessions, 2 weeks apart each. Half of these statements were true (e.g., "Kentucky was the first state west of the Alleghenies to be settled by pioneers.") and half of them were false (e.g., "Zachary Taylor was the first president to die in office."). During each session, 20 of the statements were repeated (i.e., shown at every session) and the remaining 40 were new. After the presentation phase in each session, participants rated the validity of each statement. The

authors found that participants judged repeated statements as more valid than new statements, demonstrating the basic truth effect.

Since then, a large body of research has replicated the original effect and investigated different explanations, mediators, and moderators (for a meta-analysis, see Dechêne et al., 2010; for recent summaries, Brashier & Marsh, 2020; Unkelbach et al., 2019). The effect has gained more prominence over the last years, as it may serve as an explanation for people's belief in conspiracy theories, misinformation, and fake news, due to the frequent repetition of false information on the internet and social media (Pennycook et al., 2018; Vosoughi et al., 2018). In addition, repetition trumps even knowledge about a given state of affairs (Fazio et al., 2015). However, virtually all truth effect studies relied on self-reports of subjective truth, validity, or belief, without consequences for participants. Here, we investigate what happens if a given decision (i.e., "true" or "false") has monetary consequences for the decision-maker. In other words, on a functional level, we ask if the truth effect persists if participants' decisions are (highly) incentivized.

The reasoning behind this approach is straightforward. Without consequences, participants might have little

✉ Felix Speckmann
felix.speckmann@uni-koeln.de

¹ Social Cognition Center Cologne, University of Cologne,
Richard-Strauss-Straße 2, 50931 Köln, Germany

motivation to provide correct assessments of their internal states (i.e., “Do you believe this statement?”) nor to invest too much effort into correct responses (i.e., “Is this true or false?”). In particular, when research employs online surveys, participants are likely not highly motivated to perform to the fullest of their ability. This “cognitive miser” perspective (Kurzban et al., 2013; Zipf, 1949) would predict that participants judge statements heuristically, relying on more superficial cues such as repetition and the resulting familiarity or processing fluency (see Unkelbach et al., 2019). However, if beliefs have consequences via “true/false” decisions in the form of incentives for these decisions, participants could invest more effort and potentially recall and consider more relevant knowledge that would lead to correct judgments.

Incentives as a way to increase effort are well established and can be derived from several classic theories, such as expectancy theory (Vroom, 1964), agency theory (Eisenhardt, 1989), or goal-setting theory (Locke et al., 1981). Depending on the task’s nature, the increased effort may also lead to increased performance (Bonner & Sprinkle, 2002). However, previous research has shown that the influence of incentives on several cognitive biases is small (e.g., base rate neglect, anchoring; Enke et al., 2021; Speckmann & Unkelbach, 2021). Nevertheless, observed bias reductions were mainly due to reduced reliance on intuition and reduced reliance on superficial cues, which should also reduce the effect of repetition on judged truth. Furthermore, incentives increased response times, indicating increased effort. As we use a statement set for which participants have some knowledge (Unkelbach & Stahl, 2009), increased effort implies that participants try harder and try longer to remember relevant knowledge to judge the statements.

If incentives decrease the truth effect, it would suggest that the real-life impact of repeating information is less critical than assumed so far. People likely invest some mental effort into decisions with consequences, and if such effort reduces the truth effect, it would shift the research focus on beliefs and decisions that people consider only superficially. However, if monetary incentives do not reduce the truth effect, it would underline the relevance of the phenomenon for real-life scenarios and decisions with consequences. On the theoretical level, it would show that people potentially consider repetition and its processing consequences, such as familiarity or fluency, as valid cues for their decisions (see Unkelbach & Greifeneder, 2018).

The present research

We used a standard truth effect research procedure (e.g., Bacon, 1979; Garcia-Marques et al., 2015; Unkelbach & Rom, 2017). Participants read statements in a presentation

phase, half factually true and half factually false, and a judgment phase, where participants judged in a binary-forced choice format if a given statement is “true” or “false.” Going beyond previous research, we added monetary consequences to participants’ choices: Correct responses added points and incorrect responses deducted points; these points directly translated into a monetary bonus of up to 12 Euro in a high incentives condition, 6 Euro in a medium incentives condition, and no monetary bonus in a control condition.

Given the considerations above, monetary incentives may influence the truth effect in two ways. First, participants could try to retrieve more relevant information about the presented statements. In signal detection theory terms (Swets et al., 2000), this should increase participants’ discrimination ability between factually true and factually false statements. Second, participants could try to avoid extraneous influences on their judgments, such as repetition. In signal detection theory terms, this should decrease participants’ response bias for repeated compared with new statements.

We used the statement set by Unkelbach and Stahl (2009), who showed that participants have some knowledge regarding these statements and respond more frequently “true” to repeated statements. The experiment was preregistered, and we report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures (the preregistration, data, and materials can be found at: <https://osf.io/8sj4r/>).

Method

Materials

We used 120 statements (60 true, 60 false) from Unkelbach and Stahl (2009). Table 1 shows some example statements.

Participants and design

We had no a priori estimate for the effect size of monetary incentives; we pragmatically aimed for 100 participants per condition as an established threshold in our lab (i.e., smaller effects are too costly to investigate). In the end, we recruited 321 participants on campus ($M_{\text{age}} = 23.09$ years, $SD = 6.84$; 180 female, 141 male) who participated in exchange for 4€ plus a potential bonus in the incentivized conditions. In the two incentivized conditions, participants could earn up to 12€ (high incentives condition) or 6€ (medium incentives condition), but we recruited all participants with the expectation of receiving 4€. They were randomly assigned to the high incentives, medium incentives, and no incentives conditions. There were 110 participants in the high incentive, 105 participants in the medium incentive, and 106 participants

Table 1 Examples of statements used in the experiment

Correct statements	Incorrect statements
The first windmills were built in Persia.	Henbane was a popular spice during the Middle Ages.
The cat is the only pet that does not appear in the Bible.	The world's most expensive colorant is true ultramarine.
The painting <i>Bal du moulin de la Galette</i> was painted by Renoir.	Volcanos can have a theoretical maximum elevation of about 5,000 meters.
The name of the Russian space station MIR means "peace."	Adelaide is Australia's oldest city.
Alberto Fujimori was the Japanese president from 1990 to 2000.	The world's largest lake is the Aral Sea.

in the no incentive conditions. Half of the statements were randomly sampled per participant to be shown in the presentation phase (i.e., "old" statements compared with "new" statements in the judgment phase). Half of the statements were factually true and half factually false; the other half only appeared in the judgment phase. Thus, participants judged 30 true-old, 30 false-old, 30 true-new, and 30 false-new statements.

Procedure

Experimenters approached participants on campus, led them to the laboratory, and seated them in front of a computer with a Visual Basic program already running. The program asked participants to enter their age, gender, and to indicate whether German is their native language, first foreign language, or second foreign language. The program then asked participants to turn off their cell phones to avoid cheating and explained the general setup of the experiment. Specifically, it told participants, "In the first part, you will see a list of statements. Please try to read all of the statements, even if the presentation is quick. By doing this, we want to examine certain memory processes. After that, we will continue with the judgment of statements. For each statement, please indicate by keypress whether the statement is TRUE or FALSE." In the high and medium incentives conditions, this explanation continued, "ATTENTION: During the judgment phase, you can earn up to 12€ (6€) extra. This will be explained later."

After that, the presentation phase started. To begin, participants pressed the space key, and the program stated before the presentation phase: "Please note that one-half of the statements are true and the other half is false." The program randomized statements anew for each participant; each statement appeared on-screen for 1.5 seconds with a pause between statements of 1 s. After the presentation phase, the program continued with further explanations: "We will now continue with the judgment of the statements. To this end, you will be repeatedly presented with a statement and have to decide if it is true or false. Two keys of the keyboard are marked. You can decide by using these keys. YES–TRUE: left key, NO–FALSE: right key. The key mapping will also be visible on screen." The following part dealt with the bonus payments and was only displayed to participants in

the high and medium incentives conditions: "By answering correctly or incorrectly, you can win or lose real money that will be added to your point balance. For a correct TRUE/FALSE answer, you will receive 10 (5) cents. For an incorrect answer, you will lose 10 (5) cents. You will judge 120 statements and can thus earn up to 12 (6)€! Your point balance cannot turn negative. At the end of the study, you will receive your basic compensation of 4€ on top of your point balance and see a summary of all of your answers."

After this explanation, the judgment phase began. The program asked participants to place their fingers on the marked keys ("y" and "-" on a German keyboard) and to start by pressing the space key. The judgment phase presented 120 statements, and each statement was displayed until participants pressed either one of the response keys. After the judgment phase, the program debriefed participants and showed them a summary of all questions, whether their response was correct, and how many cents (if any) they received for each question. Participants then showed the ending screen to the experimenter, who thanked participants and paid them according to their performance in the medium and high incentives conditions.

Results

Percentage of "true" judgments (PTJs)

To analyze the influence of incentives on the truth effect, we computed the percentage of "true" judgments (PTJs) for each participant by coding "true" judgments as 1 and "false" judgments as 0, and averaging the responses across the 120 decisions, separately for the four combinations of factual truth (i.e., true and false) and repetition (i.e., new vs. old). Higher PTJ values indicate a higher likelihood of responding "true" for a given statement (see Unkelbach & Rom, 2017). We then submitted the PTJs to a 2 (repetition: old vs. new) \times 2 (factual truth status: true vs. false) \times 3 (incentive: high vs. medium vs. none) mixed analysis of variance (ANOVA) with repeated measures on the first two factors. Figure 1 shows the respective means.

As Fig. 1 suggests, this analysis replicated the standard truth effect. Participants showed higher PTJs for old statements ($M = 0.633$, $SD = 0.182$) compared with new

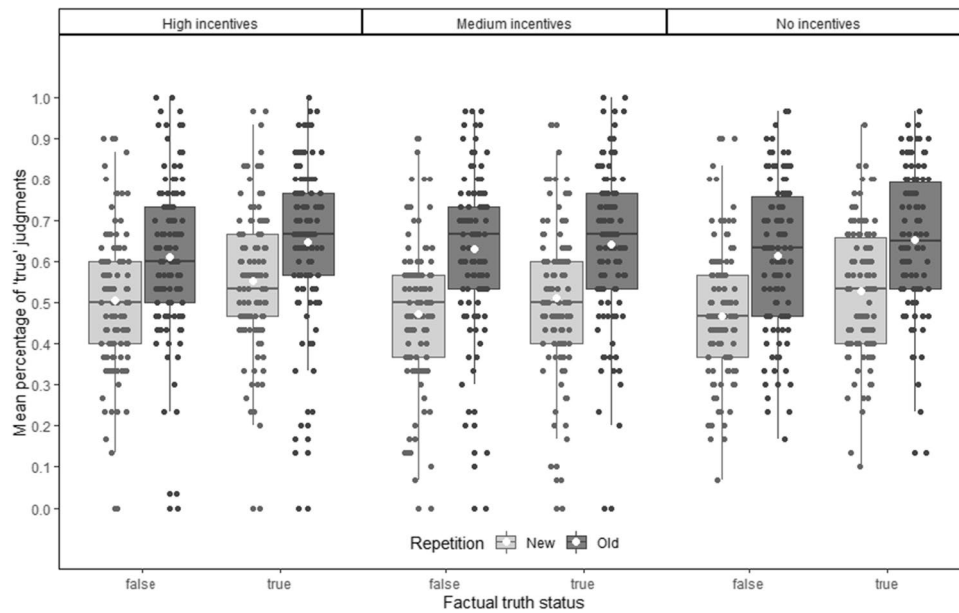


Fig. 1 Mean percentage of “true” judgments as a function of repetition (old vs. new) and factual truth status (true vs. false), separated by incentives (High vs. Medium vs. None). The white dots represent the means, the black horizontal lines represent the medians, the boxes

represent the 25% quartiles, the whiskers extend to the highest (lowest) point within the interquartile range (i.e., the distance between first and third quartile)

statements ($M = 0.507$, $SD = 0.168$), $F(1, 318) = 182.31$, $p < .001$, $\eta_p^2 = .364$. Participants also showed knowledge about the statements with higher PTJs for factually true statements ($M = 0.589$, $SD = 0.180$) compared with factually false statements ($M = 0.550$, $SD = 0.190$), $F(1, 318) = 65.00$, $p < .001$, $\eta_p^2 = .170$. However, there was no significant main effect for incentives, $M_{high} = 0.579$, $SD_{high} = 0.148$, $M_{med} = 0.564$, $SD_{med} = 0.145$, $M_{no} = 0.565$, $SD_{no} = 0.127$, $F(2, 318) = 0.40$, $p = .669$, $\eta_p^2 = .003$, and neither the knowledge effect nor the repetition-induced truth effect interacted with the incentives condition, $F(2, 318) = 2.25$, $p = .107$, $\eta_p^2 = .014$, and $F(2, 318) = 2.07$, $p = .128$, $\eta_p^2 = .013$, respectively.

In addition, the preregistered polynomial contrasts (linear and quadratic) did not interact with the repetition effect or the knowledge effect, $t(318) = 1.49$, $p = .136$, $d = 0.17$, and $t(318) = 0.21$, $p = .837$, $d = 0.02$, for the linear trends, and $t(318) = -1.37$, $p = .173$, $d = -0.15$, and $t(318) = 1.84$, $p = .066$, $d = 0.21$, for the quadratic trends.

To further explore the influence of incentives on PTJs, we also used an additional contrast testing the incentive conditions against the no incentives condition, coded -2, +1, +1, for the no, medium, and high incentive conditions, respectively. For the knowledge effect, this contrast showed no influence, $F_s < 1$. For the truth effect, however, the second contrast showed a significant effect, $t(1, 318) = 1.99$, $p = .048$, $d = 0.22$, indicating a slightly smaller truth effect in the incentive conditions ($M = 0.121$, $SD = 0.157$)

compared with the no incentive condition ($M = 0.134$, $SD = 0.187$). However, this test should be treated with caution due to the exploratory nature of this contrast and the small effect size. Furthermore, after excluding two participants who responded “false” to all questions (flat responding), the p value of this contrast changes to $p = .050$ and thus became nonsignificant by conventional standards.¹

The only other significant effect was an interaction of factual truth and repetition, $F(1, 318) = 4.75$, $p = .030$, $\eta_p^2 = .015$. The truth effect was stronger for factually false statements ($M = 0.136$, $SD = 0.190$) compared with factually true statements ($M = 0.115$, $SD = 0.184$). This effect conceptually replicates the pattern by that repetition has stronger effects on false, and thereby necessarily unknown, information (see Hasher et al., 1977; Unkelbach & Speckmann, 2021).

Because the lack of significant incentive effects on truth judgments does not provide evidence for the absence of the effect, and because the p value for the -2, 1, 1 contrast was close to the alpha level, we complemented both analyses with Bayesian analyses. As we did not preregister any priors, we used default priors for both analyses. For the main effect

¹ We repeated all analyses with the two exclusions, but none of the patterns (PTJs, correctness, or latencies) changed apart from the explorative contrast for PTJs. As we did not preregister this exclusion, we report the rest of the data without exclusions.

of incentives, we used JASP in Version 0.15 (JASP Team, 2021). A Bayesian repeated-measures ANOVA showed a Bayes factor of $BF_{01} = 20.632$, commonly seen as strong evidence for the H_0 (Jarosz & Wiley, 2014). As JASP lacks functionality for custom contrasts, we used R (R Core Team, 2020) with the packages *rstanarm* (Goodrich et al., 2020) and *bayestestR* (Makowski et al., 2019) for the exploratory $-2, 1, 1$ contrast. This analysis showed a Bayes factor of $BF_{01} = 6.410$, commonly seen as positive or substantial evidence for the H_0 (Jarosz & Wiley, 2014). The JASP file containing the analysis and the relevant R code are accessible from the OSF project linked above.

Correctness

To summarize the incentive influence on decisions, we also analyzed the effect of incentives on the overall correctness of the judgments (i.e., “true” judgment of a factually true statement or “false” judgment of a factually false statement), which provides a direct estimate of the incentive effect on decision correctness. To this end, we computed a variable indicating the correctness of each decision and submitted the average correctness, varying from 0 (*never correct*) to 1 (*always correct*) to a one-way ANOVA with incentives (high vs. medium vs. none) as the between factor. The main effect for incentives was not significant, $F(2, 318) = 2.25$, $p = .107$, $\eta_p^2 = .014$, and neither was the linear trend, $t(318) = 0.69$, $p = .489$, $d = 0.08$. However, the quadratic trend

was significant, $t(318) = 2.01$, $p = .045$, $d = 0.23$. Participants were less frequently correct in the medium incentive condition ($M = 0.513$, $SD = 0.041$), compared with the no incentives ($M = 0.525$, $SD = 0.043$) and the high incentive condition ($M = 0.521$, $SD = 0.046$). Due to the small effect size and the nonpredicted pattern, we hesitate to interpret this effect.

Latencies

We analyzed participants’ raw (i.e., no trimming or transformation) response latencies in millisecond the same way as the PTJs. Figure 2 shows the respective means. As Fig. 2 indicates, participants responded faster to old statements ($M = 3794$, $SD = 1572$) compared with new statements ($M = 4433$, $SD = 1691$), $F(1, 318) = 188.36$, $p < .001$, $\eta_p^2 = .372$. In addition, participants responded faster to factually true statements ($M = 4022$, $SD = 1556$) compared with factually false statements ($M = 4205$, $SD = 1759$), $F(1, 318) = 22.32$, $p < .001$, $\eta_p^2 = .066$, underlining that participants possessed some knowledge. In addition, incentives significantly influenced the overall latencies ($M_{high} = 4236$, $SD_{high} = 1595$; $M_{med} = 4312$, $SD_{med} = 1490$; $M_{no} = 3790$, $SD_{no} = 1375$), $F(2, 318) = 3.80$, $p = .023$, $\eta_p^2 = .023$.

To explore the influence of incentives on latencies, we used the same two contrasts for the PTJs, one testing a linear influence of incentives and one testing the two incentive conditions against the no incentives condition. Only the

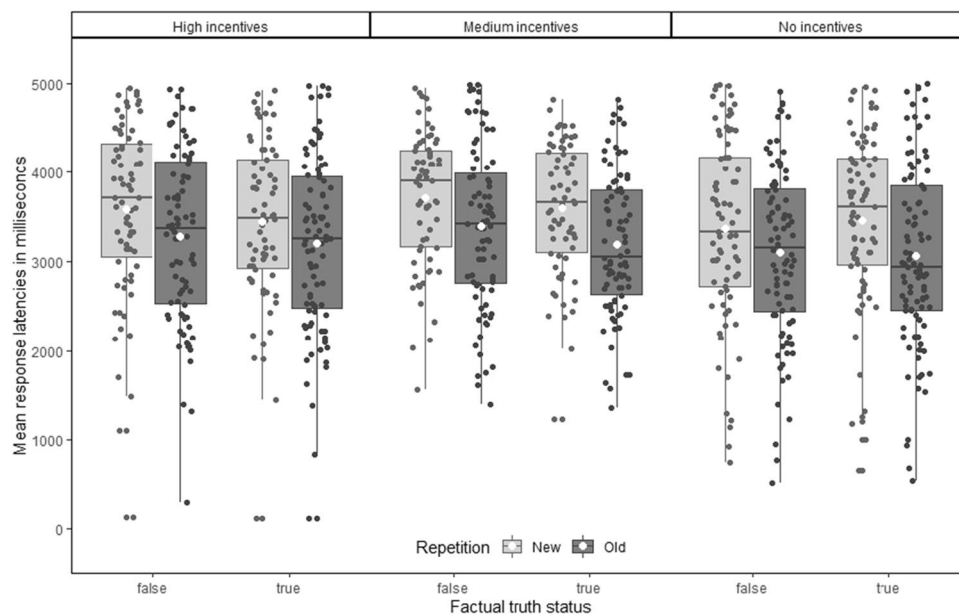


Fig. 2 Mean response latencies as a function of repetition (old vs. new) and factual truth status (true vs. false), separated by incentives (High vs. Medium vs. None). Error bars represent standard errors of the means. The white dots represent the means, the black horizontal

lines represent the medians, the boxes represent the 25% quartiles, the whiskers extend to the highest (lowest) point within the interquartile range (distance between first and third quartile)

linear contrast showed a significant effect, $t(318) = 2.20$, $p = .029$, $d = 0.25$, indicating that participants took on average more time for their true–false decisions in the high incentives condition ($M = 4236$, $SD = 1738$) compared with the no incentives condition ($M = 3789$, $SD = 1484$). Again, these contrasts are post hoc and should not be treated as confirmatory evidence.

SDT analyses

As preregistered, we also analyzed the response rates with a signal-detection theory (SDT) analysis (see Unkelbach, 2006, 2007). The SDT analysis is particularly suited for the present task, as it delivers two parameters, d' and β , which are directly interpretable as knowledge and the truth effect, respectively, in the present design. An interaction with incentives may indicate an influence of incentive on participants' higher reliance on knowledge or avoidance of bias (i.e., the truth effect). However, d' and β did not significantly differ as a function of incentives, replicating the PTJ analyses. For the complete analysis, please refer to the supplemental materials on OSF (<https://osf.io/x8wf2/>).

Discussion

The present study investigated the influence of true–false judgments' monetary consequences in a repetition-induced truth paradigm. We speculated that the monetary consequences might increase discriminability or reduce bias, thereby reducing the influence of repetition on judged truth. We replicated a typical truth effect and also the knowledge effect by Unkelbach and Stahl (2009). However, although participants could receive a bonus of up to 12€ in the high incentives condition and 6€ in the medium incentives condition, these monetary incentives did not substantially influence the truth effect or the knowledge effect. Using an exploratory contrast, we found a slight difference in the truth effect between the two incentive conditions and the no-incentive condition: Participants showed a slight reduction in their tendency to judge repeated information as true. Given the small effect size and the fact that this contrast was not preregistered, it should not be interpreted as strong evidence. If there is an effect of incentives on responses in the truth effect paradigm, it is likely minimal.

Despite the overall nonsignificant influences of incentives on the truth effect, it seems that our manipulation had the intended effect. The significant differences in response times between different incentive levels suggest that participants were more motivated to respond correctly as much as possible and consequently spent more time judging the statements.

However, it is worth noting that timing may play a part in the present patterns. Research by Jalbert et al. (2020) showed that warning participants that half of the statements are false reduces the truth effect in a truth-by-repetition paradigm. However, the warning is only effective if shown prior to exposure and ineffective when shown only before the judgment phase (see also Brashier et al., 2020). Similarly, Lewandowsky et al. (2012) recommend inserting warnings prior to the presentation phase to help people resist misinformation.

In our experiment, participants knew that their performance would be incentivized and the maximum amount of money they could earn (i.e., in the incentives conditions), but they did not concretely know how responding would be incentivized (i.e., how much money could be gained or lost for each question). This unclarity suggests the possibility that incentives might be more effective if explained in more detail prior to the presentation phase. However, before the presentation phase, participants knew that half of the statements would be false, as in the effective warning condition by Jalbert et al. (2020). Nevertheless, within the present setup, incentives neither increased participants' retrieval of relevant material from memory nor decreased their reliance on repetition as a cue for truth.

These results thus further illustrate the robustness of the truth effect by showing that even adding direct consequences to people's truth judgments does not affect it. These results are relevant as one may argue that the truth effect is often investigated with online samples of participants who might not care about their judgments because high or low performance is inconsequential. However, our data shows that the truth effect persists even when incentivizing laboratory participants with considerable amounts of money, ruling out this explanation. Our results also dovetail with a preprint manuscript by Brashier and Rand (2021); they also used a truth-by-repetition paradigm and found no effect of incentivizing a single random trial within 16 truth judgments on the truth effect, despite the fact that they provided repeated reminders about the potential reward.

Our data also fits well with existing research on other cognitive illusions, showing that incentives increase effort but not performance (Enke et al., 2021) and cognitive explanations of the truth effect. For example, the processing fluency explanation suggests that repeatedly seeing a piece of information makes it easier to process. This experienced processing fluency then serves as a cue to judge a piece of information as more true (Begg et al., 1992; Reber & Schwarz, 1999). While fluency may often be an ecologically valid cue for trueness (Reber & Unkelbach, 2010), people can also learn to use fluency as a cue for falseness (Corneille et al., 2020; Unkelbach, 2007). However, participants in the present experiment had no reason to doubt the ecological validity of fluency as a cue for truth, and thus effort did

not decrease their reliance on fluency. In terms of an incentives–effort–performance link (Bonner & Sprinkle, 2002), we provide evidence for the incentives–effort link, but the effort–performance link is disrupted, possibly due to the truth effect’s nature.

Thus, our data support existing cognitive explanations of the truth effect with potential implications for real-world phenomena (e.g., fake news, conspiracy theories, strategic misinformation): Even people who should be motivated to judge truth correctly still fall prey to the truth effect.

Code availability All code used to produce the results of the current study is available on the OSF (<https://osf.io/8sj4r>).

Authors’ contributions Christian Unkelbach (C.U.) developed the study design and wrote the computer program for data collection. Felix Speckmann (F.S.) supervised the data collection, analyzed the data, and wrote the manuscript’s largest part. C.U. wrote a smaller part and provided feedback on the overall manuscript.

Funding Open Access funding enabled and organized by Projekt DEAL. The present research was supported by a grant from the Center for Social and Economic Behavior (C-SEB) of the University of Cologne, awarded to the second author.

Data availability The data sets generated during and/or analyzed during the current study and all materials used are available on the OSF (<https://osf.io/8sj4r>).

Declarations

Conflicts of interest There are no conflicts of interest.

Ethics approval Studies of this kind do not need ethical approval at our university.

Consent to participate Participants read an informed consent sheet displayed on the computer screen and indicated their consent before beginning the study.

Consent for publication Not applicable, as no personalized data is published as part of this article.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- 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>
- 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>
- Bonner, S. E., & Sprinkle, G. B. (2002). The effects of monetary incentives on effort and task performance: Theories, evidence, and a framework for research. *Accounting, Organizations and Society*, 27(4/5), 303–345. [https://doi.org/10.1016/S0361-3682\(01\)00052-6](https://doi.org/10.1016/S0361-3682(01)00052-6)
- Brashier, N. M., & Marsh, E. J. (2020). Judging truth. *Annual Review of Psychology*, 71(1), 499–515. <https://doi.org/10.1146/annurev-psych-010419-050807>
- Brashier, N. M., & Rand, D. G. (2021). *Illusory truth occurs even with incentives for accuracy*. PsyArXiv Preprints. <https://doi.org/10.31234/osf.io/83m9y>
- Brashier, N. M., Eliseev, E. D., & Marsh, E. J. (2020). An initial accuracy focus prevents illusory truth. *Cognition*, 194, 104054. <https://doi.org/10.1016/j.cognition.2019.104054>
- 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. <https://doi.org/10.1177/1088868309352251>
- Eisenhardt, K. M. (1989). Agency theory: An assessment and review. *Academy of Management Review*, 14(1), 57–74. <https://doi.org/10.5465/amr.1989.4279003>
- Enke, B., Gneezy, U., Hall, B., Martin, D., Nelidov, V., Offerman, T., & van de Ven, J. (2021). *Cognitive biases: Mistakes or missing stakes?* (NBER Working Paper, Issue 9). <https://doi.org/10.3386/w28650>
- 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), 993–1002. <https://doi.org/10.1037/xge0000098>
- Garcia-Marques, T., Silva, R. R., Reber, R., & Unkelbach, C. (2015). Hearing a statement now and believing the opposite later. *Journal of Experimental Social Psychology*, 56, 126–129. <https://doi.org/10.1016/j.jesp.2014.09.015>
- Goodrich, B., Gabry, J., Ali, I., & Brilleman, S. (2020). rstanarm: Bayesian applied regression modeling via Stan (Version 2.21.1) [Computer software]. Retrieved October 19, 2021, from <https://cran.r-project.org/web/packages/rstanarm/index.html>
- 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. [https://doi.org/10.1016/S0022-5371\(77\)80012-1](https://doi.org/10.1016/S0022-5371(77)80012-1)
- Jalbert, M., Newman, E., & Schwarz, N. (2020). Only half of what i’ll tell you is true: Expecting to encounter falsehoods reduces illusory truth. *Journal of Applied Research in Memory and Cognition*, 9(4), 602–613. <https://doi.org/10.1016/j.jarmac.2020.08.010>

- Jarosz, A. F., & Wiley, J. (2014). What are the odds? A practical guide to computing and reporting Bayes factors. *The Journal of Problem Solving*, 7(1), 2–9. <https://doi.org/10.7771/1932-6246.1167>
- JASP Team. (2021). JASP (Version 0.15) [Computer software]. Retrieved October 19, 2021, from <https://jasp-stats.org/>
- Kurzban, R., Duckworth, A., Kable, J. W., & Myers, J. (2013). Cost-benefit models as the next, best option for understanding subjective effort. *Behavioral and Brain Sciences*, 36(6), 707–726. <https://doi.org/10.1017/S0140525X13001532>
- Lewandowsky, S., Ecker, U. K. H., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest*, 13(3), 106–131. <https://doi.org/10.1177/1529100612451018>
- Locke, E. A., Shaw, K. N., Saari, L. M., & Latham, G. P. (1981). Goal setting and task performance: 1969–1980. *Psychological Bulletin*, 90(1), 125–152. <https://doi.org/10.1037/0033-2909.90.1.125>
- Makowski, D., Ben-Shachar, M., & Lüdtke, D. (2019). bayestestR: Describing effects and their uncertainty, existence and significance within the Bayesian framework. *Journal of Open Source Software*, 4(40), 1541. <https://doi.org/10.21105/joss.01541>
- 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://doi.org/10.1037/xge0000465>
- R Core Team. (2020). R: A language and environment for statistical computing (Version 4.0.2) [Computer software]. Retrieved August 2020, from <https://www.r-project.org/>
- 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>
- Reber, R., & Unkelbach, C. (2010). The epistemic status of processing fluency as source for judgments of truth. *Review of Philosophy and Psychology*, 1(4), 563–581. <https://doi.org/10.1007/s13164-010-0039-7>
- Speckmann, F., & Unkelbach, C. (2021). Moses, money, and multiple-choice: The Moses illusion in a multiple-choice format with high incentives. *Memory & Cognition*, 49(4), 843–862. <https://doi.org/10.3758/s13421-020-01128-z>
- Swets, J. A., Dawes, R. M., & Monahan, J. (2000). Psychological science can improve diagnostic decisions. *Psychological Science in the Public Interest*, 1(1), 1–26. <https://doi.org/10.1111/1529-1006.001>
- Unkelbach, C. (2006). The learned interpretation of cognitive fluency. *Psychological Science*, 17(4), 339–345. <https://doi.org/10.1111/j.1467-9280.2006.01708.x>
- Unkelbach, C. (2007). Reversing the truth effect: Learning the interpretation of processing fluency in judgments of truth. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(1), 219–230. <https://doi.org/10.1037/0278-7393.33.1.219>
- Unkelbach, C., & Greifeneder, R. (2018). Experiential fluency and declarative advice jointly inform judgments of truth. *Journal of Experimental Social Psychology*, 79, 78–86. <https://doi.org/10.1016/j.jesp.2018.06.010>
- 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>
- 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>
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146–1151. <https://doi.org/10.1126/science.aap9559>
- Vroom, V. (1964). *Work and motivation*. John Wiley.
- Zipf, G. K. (1949). *Human behavior and the principle of least effort*. Addison-Wesley.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.