https://doi.org/10.1521/soco.2023.41.2.193

The Role of Valence Matching in the Truth-by-Repetition Effect

Jérémy Béna[a], Adrien Mierop[a], Douglas Bancu[a], Christian Unkelbach[b], and Olivier Corneille[a]

[a]UCLouvain, Belgium

[b]Social Cognition Center Cologne, University of Cologne, Germany

Accepted at *Social Cognition*

Conflict of interest statement: The authors declare no conflict of interest.

Open practices statement: The preregistration, experiment program, data, and analyses are available at https://osf.io/sbtwx/

Funding: This research was supported by a grant awarded to Olivier Corneille by IPSY (Psychological Sciences Research Institute), UCLouvain [grant IPSY-ICP-2022-1], and by an FSR Incoming Postdoctoral Fellowship [IPSY FSR22 MOVE] awarded to Jérémy Béna.

Author Note

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

Acknowledgments

We thank Océane Defresnes for contributing to discussions on this project in the context of her research internship under the supervision of Olivier Corneille.

Abstract

People judge repeated information as truer than new information, a "truth-by-repetition" effect. Because repetition increases processing fluency which is assumed to elicit positive affect, participants may match their positive experience associated with repeated information with a positive ("True") rather than negative ("False") response. We tested this valence matching hypothesis in a preregistered experiment by manipulating the affective congruency of the response format. Specifically, in the congruent condition, participants had to select a positive (negative) picture to respond "true" ("false"). In the incongruent condition, we reversed these associations. In line with the valence matching hypothesis, the truth-by-repetition effect was larger in the congruent than incongruent condition. However, the effect was small, and Bayesian analyses were inconclusive. In addition, the truth-by-repetition effect was significant in both response format conditions. The results suggest a possible contribution of a valence matching process to the truth-by-repetition effect, but one that does not challenge extant models.

Keywords: Illusory Truth Effect; Processing fluency; Truth judgment

The Role of Valence Matching in the Truth-by-Repetition Effect

People judge repeated information as truer than new information. Several explanations of this "truth-by-repetition" effect have been proposed (for recent reviews, see Brashier & Marsh, 2020; Unkelbach et al., 2019; for meta-analysis, see Dechêne et al., 2010). A prominent explanation suggests that repetition increases processing fluency which serves as a cue for judging the truth of the repeated information: what is fluent is judged as true (e.g., Reber & Schwarz, 1999; Unkelbach, 2007). Building on this processing fluency explanation, we reasoned that the match between the valence of the truth judgment (i.e., truth is positive, falsity is negative) and the affective connotation of fluency (i.e., fluency is experienced positively) might contribute to the truth-by-repetition effect. Specifically, because repetition elicits fluency (Feustel et al., 1983) and fluency elicits positive affect (e.g., Winkielman et al., 2003), participants may be biased towards matching their judgments of repeated statements with a positive response ("Statement is true") rather than with a negative response ("Statement is False"). In a typical truth-by-repetition procedure, the truth and the valence components of the responses are confounded, and this confound may contribute to the effect.

Fluency, Valence, and the truth-by-repetition effect

The truth-by-repetition effect is most likely determined by several processes. Two processes received more attention (for overviews, see Unkelbach & Rom, 2017; Unkelbach et al., 2019). First, repetition increases subjective truth because people consciously recognize the statements (Arkes et al., 1991; Bacon, 1979; Brown & Nix, 1996). Second, besides conscious memory, the truth-by-repetition effect has an experiential component (Dechêne et al., 2010; Unkelbach & Stahl, 2009; but see Mattavelli et al., 2022). That is, repetition increases subjective truth even without conscious source recollection (Begg et al., 1992). Reber and Schwarz (1999) suggested that the experience of processing fluency increases subjective truth independent of memory by manipulating the color contrast of statements (see also Hansen et al., 2008; Silva et al., 2016; Unkelbach, 2007). Reber and Unkelbach (2010) also provided an epistemological justification for this effect: people use processing fluency as a cue for truth judgments because, ecologically, truth and fluency are correlated (see also Herzog et al., 2013; Unkelbach & Greifeneder, 2013).

Multiple fluency accounts have been offered to explain why repetition increases truth judgments. Among them, there are three main accounts. A first, ecological account, which we have just alluded to, considers fluency as an ecologically valid cue for truth in default contexts. A second, amplification account (e.g., Albrecht & Carbon, 2014, Landwehr & Eckmann, 2020), states that fluency increases judgments no matter judgment content (but see Corneille et al., 2020). A third, "hedonic marking" account, is more central to this investigation. It proposes that fluency elicits positive feelings (e.g., Reber et al., 1998; Winkielman & Cacioppo, 2001; see Winkielman et al., 2003), and that these positive feelings increase judgments on dimensions that are felt as positive, such as "true."

If truth is connotated positively (see, for example, Warriner et al., 2013[1]), then a valence matching process may also contribute to the experiential component of the truth-by-repetition effect. This conclusion follows because participants may match their affective experiences with

[1] Warriner et al. (2013) collected word norms of valence for 13,915 English lemmas, including “truth,” “true,” “untrue,” and “false.” Participants rated the pleasantness of the emotions evoked by several words on a 9-point Likert scale ranging from “unhappy” to “happy.” Overall, Warriner et al. found that the mean pleasantness was 7.19 (*SD* = 1.78) for “truth,” 7.51 (*SD* = 1.67) for “true,” 3.45 (*SD* = 1.82) for “untrue,” and 3.27 (*SD* = 2.06) for “false.” These results suggest that concepts related to truth are generally experienced as much more pleasant than concepts related to falsehood.

affectively consistent *responses* (e.g., the affective-matching model, Klauer & Musch, 2003; see also Wentura, 2000).

In the truth effect procedure, this corresponds to matching the positive experience conveyed by repetition with positively-connotated truth responses. This valence matching process would contribute to the truth-by-repetition effect because "true" responses, just as repetition-induced fluency, are positive. If supported, this valence-matching process would challenge other fluency explanations of the truth-by-repetition effect.

Unkelbach et al. (2011) and Hilbig (2009, 2012; see also Jaffé & Greifeneder, 2021) already addressed the potential effect of positivity on judged truth. However, they focused on the affective connotation of the statements rather than the response options (i.e., true vs. false); that is, they investigated if people believe positive and negative information more (for other manipulations, see also Koch & Forgas, 2012; Stump et al., 2021). Unkelbach et al. (2011) found that statements' valence did not moderate the truth-by-repetition effect (Experiments 1 & 2), but they found a positive correlation between statements' rated valence and perceived truth (Experiment 1). Hilbig (2009, 2012) reported greater perceived truth for statements that were framed negatively than positively. In contrast, we hypothesized that information might acquire its affective connotation through repetition-induced fluency, and because the valence of this positive affective experience matches the positive valence of truth judgments, a truth-by-repetition effect emerges.

The hypothesis that valence matching might contribute to the truth-by-repetition effect was suggested by the results of a recent experiment by Corneille et al. (2020, Experiment 3; for a recent replication, see Béna et al., 2022). In this experiment, participants either judged the truth ("Is this a true statement?" Yes vs. No) or falsity ("Is this a false statement?" Yes vs. No) of repeated and new statements. This procedure departed from standard truth-by-repetition studies where participants have to select between a "True" vs. "False" response option for each presented statement. Corneille et al. found a truth-by-repetition effect for both “truth” and “falsehood” judgments. Of note, however, an asymmetry of repetition effects was found. Specifically, repetition increased the proportion of “true” judgments to a larger extent than it decreased the proportion of “false” judgments. This asymmetry is in line with a valence matching hypothesis. When judging the truth of the statements, the repetition-induced fluency is positive, and the "Yes, true" judgment is also positive. In contrast, when judging the falsehood of the statements, the repetition-induced fluency remains positive, but this time, it does not match in valence with the negativity of a “No, not false” response.

However, there are other explanations for the observed asymmetry (e.g., higher difficulty in the falsity condition). Thus, in the present research, we directly tested the possible contribution of valence matching between the "true" response option and repetition-induced fluency (again: both positive) to the truth-by-repetition effect by designing a variant of a study by Corneille et al. (2005). Specifically, in a congruent condition, participants selected the "true" response option by using a positive response format (e.g., the picture of a butterfly), and they selected the "false" response option by using a negative response format (e.g., the picture of a snake), and vice-versa in the response incongruent condition. Thus, in one condition, the responses' evaluative connotation matched with fluency's evaluative connotation but mismatched in the other condition.

We predicted to replicate the truth-by-repetition effect in the congruent condition: repeated statements should be more likely to be classified as "true" relative to new statements when the "true" response option is associated with a positive picture. The critical question, however, is whether and to what extent the truth-by-repetition effect would be reduced in the incongruent condition (when the "true" response option is associated with a negative picture). If people use the evaluative connotation of processing fluency to judge a statement as "true," then the truth-by-repetition effect should be significantly reduced in the incongruent condition. In the incongruent condition, participants can no longer match their positive feeling associated with repetition-induced fluency with a truth response, as truth is now associated with a negative picture. How large this reduction is (possibly, reducing the truth-by-repetition effect to non-significance) indicates the contribution of a valence matching process to the truth-by-repetition effect.

Three comments are in order before proceeding to the presentation of the experiment. First, the present valence matching hypothesis is consistent with a fluency account. To illustrate this, let us assume a radical valence matching account, in which no other process than valence matching influences truth judgments. The only assumptions would be that fluency is positively connotated (see Winkielman et al., 2003) and that "truth" is more positive than "falsity" (see Warriner et al., 2013). There is no need for ecological correlations (e.g., Reber & Unkelbach, 2010) or for fluency/familiarity to be (mis)attributed to truth (Begg et al., 1992). Second, a valence matching account does not need to be radical. As the truth-by-repetition effect is most likely not process-pure, valence matching may contribute, to some extent, to the truth-by-repetition effect in usual response valence conditions, while other processes play a role as well. Third, to test the valence matching hypothesis, truth needs not to be positively connotated in all possible contexts (it is conceivable that, in some situations, truth is negatively connotated); all that is needed is a general context (as the one implemented in the study we report below) where one can assume that the concept of "truth" evokes more positive feelings than the concept of falsehood (see Warriner et al., 2013; see Footnote 1 above).

Methods

We report how we determined our sample size, all data exclusions, all manipulations, and all measures. The preregistration, experiment program, data, and analysis scripts are publicly available (Béna et al., 2022) at https://osf.io/sbtwx/.

We implemented a 7-day time interval between the exposure phase and the truth judgment task. This time interval should increase the relative contribution of experiential processes in the truth-by-repetition effect as it decreases the referential component of the truth-by-repetition effect. From a recognition memory perspective, introducing a time interval reduces conscious recollection (see, e.g., Yonelinas, 2002). As a result, introducing a time interval allows for a strong test of the present valence matching hypothesis.

Design

The experiment used a 2 (Response format: Congruent vs. Incongruent) × 2 (Repetition status: Repeated vs. New) mixed design, with Response format manipulated between participants and Repetition status manipulated within participants. Seven to eight days after the exposure phase, participants were randomly allocated to the Congruent condition or the Incongruent condition. In the Congruent condition, participants had to use positive pictures to indicate a "true" response and negative pictures to indicate a "false" response. Conversely, in the Incongruent condition, participants had to use negative pictures to indicate a "true" response and positive pictures to indicate a "false" response.

Participants

Based on resource constraints, we preregistered a sample size of at least 200 participants. We collected the data online on Prolific. In the first session, we collected data from 335 participants because we expected a dropout between the exposure phase and the truth judgment task one week after about 40%. Out of the 335 participants, 315 provided complete data. In the second session, 259 participants returned, of which 249 provided complete data (50.2% female; 49.8% male; M_{age} = 41.53; SD_{age} = 14.42, 1 not reported).

Participants were English speakers, had an approval rate of at least 95%, had at least 100 previous submissions, and did not take part in two similar experiments (and pre-tests) that we conducted before this experiment[2]. Participants were paid US $ 1.68 for completing both sessions (estimated total completion time: 10 minutes). Based on the preregistered criteria, we excluded 77 of the 315 participants that completed the first session (66 did not come back to the truth judgment task or did not provide data; 11 declared they did not pay attention to the stimuli or that they did not take their responses seriously). This resulted in a final sample size of 238 participants.

Sensitivity power analyses (performed with G*Power 3.1.9.7, Faul et al., 2007) showed that N = 238 yields a statistical power of $1 - \beta = .80$ (.90) to detect an effect of Response Format on truth-by-repetition effect scores (i.e., the difference between proportions of "true" judgments

[2] We cannot report these initial experiments as they used IAPS pictures, which, unknown to us at the time we collected the data, are not allowed in online research. These initial experiments were very similar to the present experiment (the second one included a 7-day time interval between the exposure and truth judgment phases; the first one did not include such a time interval). In the two studies, the critical interaction between Repetition and Matching was not statistically significant. These findings are consistent with the caution indicated in the abstract and General Discussion when interpreting the interaction we found in the present experiment.

on repeated vs. new judgments, capturing the truth-by-repetition effect) as small as Cohen's d = 0.325 (d = 0.382 – small-to-medium effects) in a one-tailed independent samples t-test (α = .05).

Materials and procedure

We based our materials and procedure on Unkelbach and Rom (2017). Similar to Corneille et al. (2005), we allocated participants randomly to a Congruent (n = 107) or Incongruent (n = 131) Response format (see below). We conducted the experiment online with lab.js (Henninger et al., 2021), JATOS (Lange et al., 2015), and Prolific. First, participants read a consent form informing them that the experiment was about various influences on people's assessment of statements.

The presentation phase involved 30 statements randomly selected from a pool of 60 statements (30 true; 30 false) and sequentially presented onscreen for 2.5s each, with a 1s break between statements. The statements were emotionally neutral; 15 were factually true, and 15 were false. Participants were asked to read the statements. After the exposure phase, we administered an attention check in which we asked participants whether they read all the displayed statements. We excluded participants who failed this check (see the exclusion criteria above).

Participants were invited to complete the second session seven to eight days after the first session. Participants had to judge 60 statements (30 repeated; 30 new) as "True" or "False." Figure 1 displays the truth judgment's trial structure with an example of a statement we used. The question "Is this statement true or false?" was visible on the upper half of the screen. Statements remained visible onscreen until participants responded by clicking on the positive [negative] picture associated with the "True" response or on the negative [positive] picture associated with the "False" response in the Congruent [Incongruent] condition. Pictures were

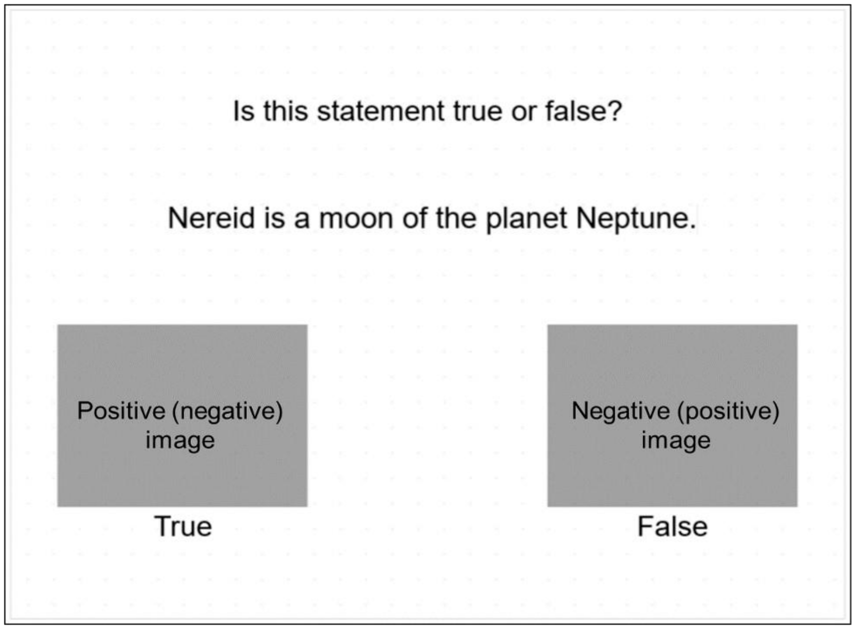

Figure 1. Trial structure (with a statement we used as an illustration) of the truth judgment task. In the Congruent condition, a positive picture was associated with a "True" response" and a negative picture was associated with a "False" response, while the opposite was the case in the Incongruent condition. Pictures were selected from the OASIS (Kurdi, Lozano, & Banaji, 2017).

always displayed above the "True"/"False" labels of the response options they corresponded to, as displayed in Figure 1. The "True" response option and its accompanying picture were always displayed on the left side of the screen; the "False" response option with its accompanying picture were on the right side of the screen. We used 30 positive and 30 negative pictures from the Open Affective Standardized Image Set (OASIS; Kurdi, Lozano, & Banaji, 2017; for the pictures selection, see https://osf.io/65u8m). Each pair of pictures (one positive, one negative) was displayed twice in a random order with a random statement.

Once a response was entered, the statement and the pair of pictures were removed from the screen and replaced by a 1s blank screen, directly followed by a new trial. Participants were reminded of the instructions (i.e., which picture should be selected to signal truth and falsehood) at the bottom of each screen. This reminder, in conjunction with the manipulation of Response format as a between-participants factor, made it unlikely that participants could be confused using the response keys.

After the truth judgment task, we administered a seriousness check (based on Aust et al., 2013) and excluded participants who failed it (see above). We also asked participants whether they searched for information about the statements (during the study or between sessions). We conducted the preregistered and non-preregistered analyses with and without participants responding that they did so, without consequences on our conclusions. Participants then indicated what they thought the experiment's goal was and any comments they had. Participants were then thanked and debriefed.

Results

We computed proportions of "true" judgments (the number of "true" responses divided by the number of statements) at each level of Repetition × Response format condition for each participant. Proportions of "true" judgments are typically computed to analyze the data of truth-by-repetition studies (e.g., Unkelbach & Rom, 2017). Figure 2 displays the Proportions of "true" Judgments as a function of Repetition and Response format. The valence matching hypothesis predicts that the truth-by-repetition effect (i.e., as estimated by the difference between proportions of "true" judgments on repeated and new statements, "truth-by-repetition effect scores" hereafter) should be larger in the Congruent condition than in the Incongruent condition. The critical test we preregistered was a one-tailed independent samples *t*-test with truth-by-repetition effect scores as the dependent variable and Response format as the independent variable; the truth-by-repetition effect is the within-participants difference in proportions of "true" judgments on repeated vs. new statements. We performed the analyses in R (R Core Team, 2021) and the one-tailed independent samples *t*-test in base R.

We also conducted a non-preregistered one-sided Bayesian independent samples *t*-test (with the 'BayesFactor' package, Morey et al., 2018; we used a Cauchy distribution with √{2}/2 as the scale value). We tested the hypothesis that the valence matching effect is positive (i.e., truth-by-repetition scores Congruent condition – truth-by-repetition scores Incongruent condition > 0) against the point null (i.e., truth-by-repetition scores Congruent condition – truth-by-repetition scores Incongruent condition = 0). We used conventional cut-offs to interpret the Bayes Factors (BF; e.g., Dienes, 2014). A BF above 3 yields evidence for the alternative hypothesis (here, the valence matching hypothesis) compared to the null hypothesis, while a BF below 1/3 yields evidence for the null hypothesis. A BF between 1/3 and 3 would count as anecdotal evidence in either direction, and indicates insensitive data (i.e., no evidence for the hypotheses).

In addition, we analyzed the proportions of "true" judgments in a preregistered frequentist 2 (Response format: Congruent vs. Incongruent) × 2 (Repetition: Repetition vs. New) mixed ANOVA with repeated measures on the latter factor (with the 'afex' package, Singmann et al., 2021). We also conducted a preregistered Bayesian mixed ANOVA (with the 'BayesFactor' package). We used the default medium r scale (1/2) for both fixed (Repetition status, Response format, and their interaction) and random effects (participants). As we were interested in Bayes Factors (BF) for the main effects and the two-way interaction between Repetition status and Response format, we computed the Bayes Factor of each model against the null hypothesis (of no effect) and contrasted the performance of the models of interest to the relevant model omitting the effect. For example, a BF of 4 indicates that data are four times

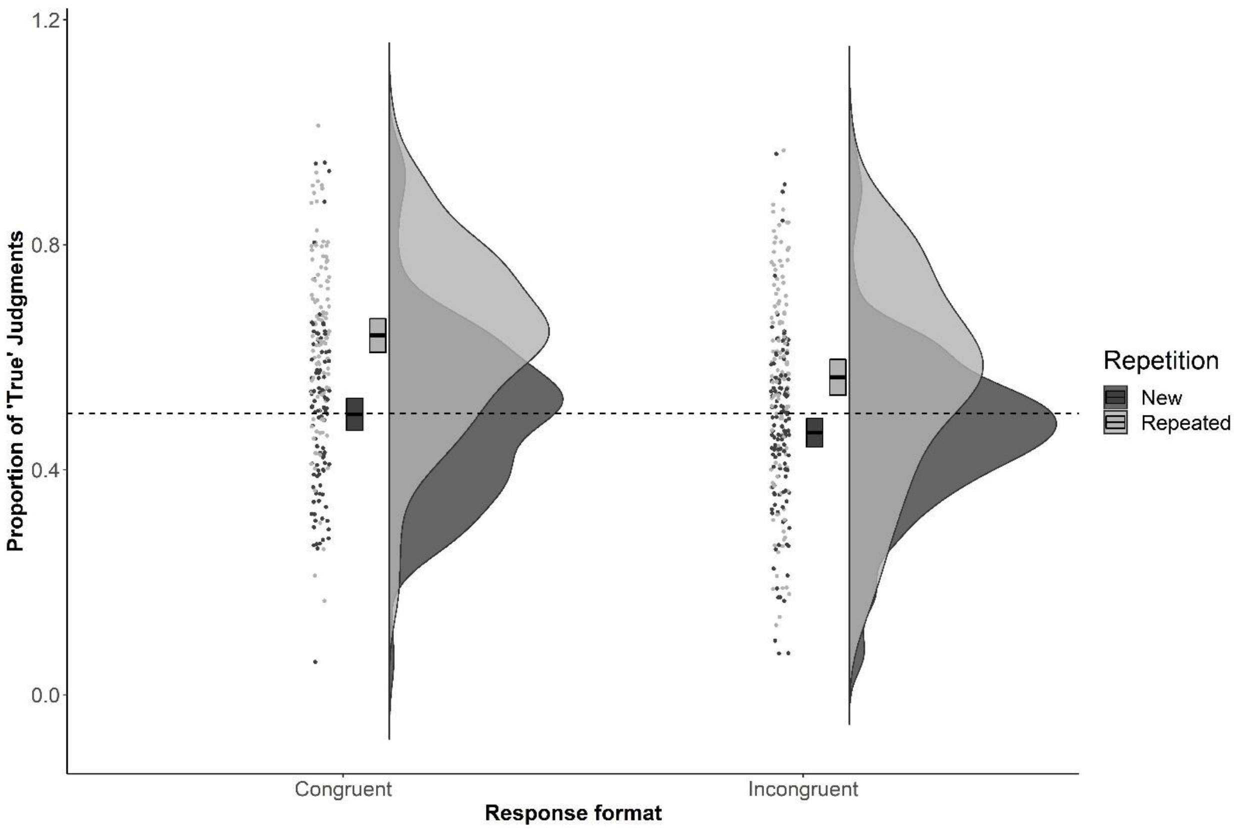

Figure 2. Proportion of "true" judgments as a function of Repetition status (Repetition vs. New) and Response format condition (Congruent vs. Incongruent). The dots are the participants' scores (jittered). The lower and upper limits of the boxplots are the 95% confidence intervals, with the mean in between. The distributions represent the kernel probability density of the data. The dashed line represents the chance level (50% of "true" judgments).

more likely under the model with the relevant effect than under the model without it. A BF of .25 means that data are four times more likely under the model without the relevant effect than under the model with it.

In the critical one-tailed independent samples *t*-test, we found that truth-by-repetition effect scores were significantly higher in the Congruent condition (M = 0.14; SD = 0.19) than in the Incongruent condition (M = 0.10; SD = 0.18), $t(236) = 1.79$, $p = .038$, $d = 0.23$. The one-sided Bayesian paired-samples *t*-test, however, yielded inconclusive evidence, as the BF was very close to 1, BF = 1.22.

Additional non-preregistered one-sample *t*-tests (test value = 0) in each Response format condition showed that the truth-by-repetition effect was significant both in the Congruent condition, $t(106) = 7.80$, $p < .001$, $d = 0.75$, and in the Incongruent condition, $t(130) = 6.41$, $p < .001$, $d = 0.56$.

We now turn to preregistered complementary analyses to investigate main effects we could not test in the critical *t*-test reported above[3]. In the mixed ANOVA[4], we observed a strong

---

[3] The interaction between Repetition and Matching condition is already being tested in the critical preregistered one-sided independent samples *t*-test; as noted in the main text, the mixed ANOVA was mainly aimed at testing the main effects of Repetition and Matching. Still, we report the interaction between Repetition and Matching for completeness, but note that the statistical power of this test is lower than in the *t*-test (the *t*-test is one-sided; the mixed ANOVA is two-sided as it cannot be one-sided). Based on the means and standard deviations we observed at each Repetition × Matching level, we simulated power with the R package Superpower (Lakens & Caldwell, 2021). To achieve enough power (80%) to detect an interaction between Repetition and Matching based on the pattern of results we observed, we would need $n = 203$ in each Matching condition, 406 participants total – which was well beyond our resources. With $n = 107$ (131) (the sample size we have in the congruent [incongruent] condition), power is about 55.2% (62.25%).

[4] A reviewer suggested that we also conduct a non-preregistered generalized linear mixed-effects (GLMM) model to analyze the categorical responses without first aggregating responses across statements. We modeled the categorical dependent variable (“true,” “false”) as a function of three fixed effects (Repetition and Matching—both mean-centered, Sommet & Morselli, 2017, and their interaction) and two random intercepts (participants and statements). We tested a logistic mixed-effects regression. We used the R packages ‘lme4’ (Bates et al., 2015; version 1.1-27.1), ‘sjPlot’ (Lüdecke, 2021; version 2.8.10), and ‘ggeffects’ (Lüdecke, 2018; version 1.1.1). We report the odds ratios (OR) and their 95% confidence intervals (CI) for each effect below. When CI do not include 1, we interpret the effect as significant; if they include 1, we interpret the effect as not significant. We found a significant main effect of Repetition status, OR = 1.73, 95%CI = [1.61, 1.86]: Truth judgments were higher for repeated ($M = 0.61$; 95%CI = [0.57, 0.65]) than new ($M = 0.48$; 95%CI = [0.44, 0.52]) statements (i.e., the truth-by-repetition effect). The effect of Response format was also significant, OR = 1.28, 95%CI = [1.09, 1.50]: Truth judgments were higher in the Congruent ($M = 0.58$; 95%CI = [0.54, 0.62]) than Incongruent ($M = 0.52$; 95%CI = [0.48, 0.56]) condition. Critically, the interaction between Repetition status and Response format was significant, OR = 1.19, 95%CI = [1.04, 1.38]. In line with the critical test we preregistered (the one-sided independent samples *t*-test), the difference between repeated and new statements was larger in the Congruent ($M_{new} = 0.5$; 95%CI = [0.45, 0.55]; ($M_{rep} = 0.65$; 95%CI = [0.61, 0.70]) than Incongruent ($M_{new} = 0.46$; 95%CI = [0.42, 0.50]; ($M_{rep} = 0.58$; 95%CI = [0.53, 0.62]) condition.

main effect of repetition, $F(1, 236) = 103.18$, $p < .001$, $\eta^2_p = .304$, which was also supported by the Bayesian analysis, BF = 7.61e16 ± 2.48%: participants judged repeated statements more often as true ($M = 0.60$, $SD = .17$) than new statements ($M = .48$, $SD = .15$). The main effect of Response format was also significant, $F(1, 236) = 9.62$, $p = .002$, $\eta^2_p = .039$, which was also supported by the Bayesian analysis, BF = 12.15 ± 2.75%: participants in the Congruent condition judged statements more often as true ($M = 0.57$, $SD = .12$) than participants in the Incongruent condition ($M = .52$, $SD = .14$). The test of the interaction between Repetition and Response format, which corresponds to a non-directional - and so less powerful - test for the valence matching effect addressed by the aforementioned one-sided *t*-test, was not significant, $F(1, 236) = 3.19$, $p = .075$, $\eta^2_p = .013$, and the Bayesian analysis yielded no conclusive evidence, BF = 0.62 ± 2.54%.

We also preregistered analyses retaining the 233 (out of 238) participants declaring having not searched for any information about the statements during the study or the time interval. These analyses yielded similar results[5].

General Discussion

People believe repeated information more than new information. This truth-by-repetition effect is commonly explained with fluency accounts. However, because feelings of fluency and

---

[5] The one-tailed independent samples *t*-test again showed a significant effect of Response format on truth-by-repetition effect scores, $t(231) = 1.81$, $p = .036$, $d = 0.24$, BF = 1.28. In the 2 (Response format: Congruent vs. Incongruent) × 2 (Repetition: Repeated vs. New) mixed ANOVA, we found the expected truth-by-repetition effect, $F(1, 231) = 103.53$, $p < .001$, $\eta^2_p = .309$, which was also supported by the Bayesian analysis, BF = 9.43e16 ± 4.51%. We also observed a main effect of Response format, $F(1, 231) = 10.58$, $p = .001$, $\eta^2_p = .044$, and the Bayesian analysis yielded evidence for it, BF = 20.24 ± 4.75%. The interaction between Repetition and Response format was not significant, $F(1, 231) = 3.26$, $p = .072$, $\eta^2_p = .014$, and the Bayesian analysis was inconclusive, BF = 0.64 ± 5.62%.

the notion of truth are both positively connotated, we reasoned that the truth-by-repetition effect might capture the expression of a positive judgment if participants match their affect with truth responses. To examine this possibility, we compared conditions in which the valence of the response format matched or mismatched the truth judgments' valence. Overall, we observed a strong truth-by-repetition effect. For the main test, we preregistered a one-sided *t*-test. In addition, we conducted a non-preregistered logistic mixed-effects regression. We observed that the truth-by-repetition effect was qualified by the response format in our main test: When the response format was affectively incongruent with the truth response (i.e., "true" was associated with a negative picture and "false" with a positive picture), the truth-by-repetition effect was smaller but still statistically significant and medium-sized than when the response format was consistent with the truth responses (i.e., "true" was associated with a positive picture and "false" with a negative picture). Neither the frequentist ANOVA (although it was underpowered for this aim, see Footnote 3) nor the Bayesian analyses, however, yielded support for a valence matching contribution to the truth-by-repetition effect. Thus, one should remain cautious about conclusively stating the existence of a valence matching contribution to the truth-by-repetition effect. The contribution of valence matching to the truth-by-repetition effect may be considered small and secondary, even under the optimal conditions examined here (that involved a one-week delay aimed at weakening conscious recollection).

Response-matching-driven biases pose serious challenges for many models (for an overview, see Proctor & Cho, 2006). For instance, Fiedler and colleagues (2001) observed that mood-congruent memory (i.e., enhanced memory for stimuli that match an individual's emotional state) is partly explained by the response bias to report memories that match one's current mood. The present research clarifies that valence matching influence is likely to play only a minor role in the truth-by-repetition effect, if any. As a matter of fact, we found at least a medium truth-by-repetition effect in both the Congruent and Incongruent conditions. Hence, fluency models that do not primarily depend on processing fluency's evaluative connotation are clearly relevant when accounting for the truth-by-repetition effect; for example, that fluency is a cue that is ecologically correlated with truth (e.g., Reber & Unkelbach, 2010; Unkelbach & Greifeneder, 2013) or that people use heuristics that translate fluency experiences into observable judgments (e.g., Schwarz, 2004; Hertwig et al., 2008).

Finally, we find it relevant to briefly discuss the present results in light of the asymmetry Corneille et al. (2020, Experiment 3) reported: The authors found a smaller truth-by-repetition effect for truth than falsehood judgments. When a "falsehood" scale is used, participants cannot match the positive experience elicited by the repetition-induced fluency with a "true" response, as none is offered on such a scale. In contrast, when a "truth" scale is used, participants can match their positive experience with a "true" response. Considering the current findings, we think it unlikely that valence matching fully explains this asymmetry. In the present experiment, we blatantly paired truth response options with positive or negative pictures, and we additionally introduced a one-week delay to increase the experiential component of the truth-by-repetition effect. Even under such conditions, the valence matching effect was small.

The two fluency accounts we mentioned above (i.e., the ecological account and the naïve theories account) could be used to explain the smaller truth-by-repetition effect observed for falsehood than truth judgments. First, in line with Reber and Unkelbach (2010), participants may learn that fluency is a cue that predicts truth, but not that the cue predicts the absence of falseness. From a learning perspective (Unkelbach, 2006), it is possible that people more readily learn the correlation between present states (i.e., "fluency" and "truth") rather than the correlation between a present and an absent state (i.e., "fluency" and "not false"). Second, people may have the naïve theory or judgmental heuristic that fluent processing indicates truth. It may be quite a different heuristic that fluent processing indicates "not false." Given that valence matching does not appear to play a major role in the occurrence of the truth-by-repetition effect, the non-fit of the learned correlation or the non-applicability of the heuristic may explain the asymmetrical pattern obtained by Corneille et al., 2020).

Conclusion

The truth-by-repetition effect is most likely determined by multiple factors. As processing fluency is a prominent factor, we hypothesized that valence matching of processing fluency's positive connotation with the positive connotation of the "true" response might partially explain the impact of the fluency factor. By manipulating the response format, we could isolate the potential contribution of valence matching. We found a strong truth-by-repetition effect, which proved larger in the congruent than incongruent condition in our main analysis. Overall, our preregistered and non-preregistered analyses indicated that the valence matching effect might be considered secondary because the influence of valence matching in the truth-by-repetition effect was minor even in conditions meant to promote it (i.e., when using a one-week delay aimed at weakening conscious recollection). Other theoretical accounts, and in particular cue learning or naïve metacognitive theories, seem better suited to explain why repetition causes people to believe information more.

References

Albrecht, S., & Carbon, C. C. (2014). The fluency amplification model: Fluent stimuli show more intense but not evidently more positive evaluations. *Acta Psychologica, 148,* 195-203. https://doi.org/10.1016/j.actpsy.2014.02.002

Arkes, H. R., Boehm, L. E., & Xu, G. (1991). Determinants of judged validity. *Journal of Experimental Social Psychology*, *27*(6), 576-605. https://doi.org/10.1016/0022-1031(91)90026-3

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

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1-48. https://doi.org/10.18637/jss.v067.i01

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

Béna, J., Bancu, D., Unkelbach, C., & Corneille, O. (2022). Valence matching in the repetition-induced truth effect: An experiment with a 7-day delay between exposure and test using valenced images from the Open Affective Standardized Image Set. OSF Repository. https://doi.org/10.17605/OSF.IO/SBTWX

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. http://doi.org/10.5334/irsp.683

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

Brown, A. S., & Nix, L. A. (1996). Turning lies into truths: Referential validation of falsehoods. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(5), 1088-1100. https://doi.org/10.1037/0278-7393.22.5.1088

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

Corneille, O., Monin, B., & Pleyers, G. (2005). Is positivity a cue or a response option? Warm glow vs evaluative matching in the familiarity for attractive and not-so attractive faces. *Journal of Experimental Social Psychology*, *41*(4), 431-437. https://doi.org/10.1016/j.jesp.2004.08.004

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00781

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

Feustel, T. C., Shiffrin, R. M., & Salasoo, A. (1983). Episodic and lexical contributions to the repetition effect in word identification. *Journal of Experimental Psychology: General*, *112*(3), 309. https://doi.org/10.1037//0096-3445.112.3.309

Fiedler, K., Nickel, S., Muehlfriedel, T., & Unkelbach, C. (2001). Is mood congruency an effect of genuine memory or response bias? *Journal of Experimental social psychology*, *37*(3), 201-214. https://doi.org/10.1006/jesp.2000.1442

Hansen, J., Dechêne, A., & Wänke, M. (2008). Discrepant fluency increases subjective truth. *Journal of Experimental Social Psychology*, *44*(3), 687-691. https://doi.org/10.1016/j.jesp.2007.04.005

Henninger, F., Shevchenko, Y., Mertens, U. K., Kieslich, P. J., & Hilbig, B. E. (2021). lab.js: A free, open, online study builder. *Behavior Research Methods.* https://doi.org/10.3758/s13428-019-01283-5

Hertwig, R., Herzog, S. M., Schooler, L. J., & Reimer, T. (2008). Fluency heuristic: A model of how the mind exploits a by-product of information retrieval. *Journal of Experimental Psychology: Learning, memory, and cognition, 34*, 1191-1206. https://doi.org/10.1037/a0013025

Herzog, R., Álvarez-Pasquin, M. J., Díaz, C., Del Barrio, J. L., Estrada, J. M., & Gil, Á. (2013). Are healthcare workers' intentions to vaccinate related to their knowledge, beliefs and attitudes? A systematic review. *BMC public health, 13*(1), 154. https://doi.org/10.1186/1471-2458-13-154

Hilbig, B. E. (2009). Sad, thus true: Negativity bias in judgments of truth. *Journal of Experimental Social Psychology*, *45*(4), 983-986. https://doi.org/10.1016/j.jesp.2009.04.012

Hilbig, B. E. (2012). How framing statistical statements affects subjective veracity: Validation and application of a multinomial model for judgments of truth. *Cognition, 125*(1), 37–48. https://10.1016/j.cognition.2012.06.009

Jaffé, M. E., & Greifeneder, R. (2021). Negative or Negated, Thus True? An Investigation of Concept Valence and Semantic Negation as Drivers of Framing Effects in Judgments of Truth. *Social Cognition*, *39*(6), 687–716. https://doi.org/10.1521/soco.2021.39.6.687

Koch, A. S., & Forgas, J. P. (2012). Feeling good and feeling truth: The interactive effects of mood and processing fluency on truth judgments. *Journal of Experimental Social Psychology*, *48*(2), 481–485. https://doi.org/10.1016/j.jesp.2011.10.006

Klauer, K. C., & Musch, J. (2003). Affective priming: Findings and Theories. In J. Musch & K. C. Klauer (Eds.), *The Psychology of Evaluation: Affective Processes in Cognition and Emotion* (1st ed.) (pp. 7-36). Psychology Press.

Kurdi, B., Lozano, S., & Banaji, M. R. (2017). Introducing the Open Affective Standardized Image Set (OASIS). *Behavior Research Methods, 49*(2), 457–470. https://doi.org/10.3758/s13428-016-0715-3

Lakens, D., & Caldwell, A. R. (2021). Simulation-Based Power Analysis for Factorial Analysis of Variance Designs. *Advances in Methods and Practices in Psychological Science*, *4*(1). https://doi.org/10.1177/2515245920951503

Landwehr, J. R., & Eckmann, L. (2020). The nature of processing fluency: Amplification versus hedonic marking. *Journal of Experimental Social Psychology, 90,* 103997. https://doi.org/10.1016/j.jesp.2020.103997

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1999). International affective picture system (IAPS): Instruction manual and affective ratings. *The center for research in psychophysiology, University of Florida.*

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

Lüdecke D (2018). "ggeffects: Tidy Data Frames of Marginal Effects from Regression Models." *Journal of Open Source Software, 3*(26), 772. https://doi.org/10.21105/joss.00772

Lüdecke D (2021). sjPlot: Data Visualization for Statistics in Social Science. R package version 2.8.10. URL: https://CRAN.R-project.org/package=sjPlot.

Mattavelli, S., Corneille, O., & Unkelbach, C. (2022). Truth by Repetition… without repetition: Testing the effect of instructed repetition on truth judgments. In press at *Journal of Experimental Psychology: Learning, Memory & Cognition*. https://doi.org/10.1037/xlm0001170

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

Proctor, R. W., & Cho, Y. S. (2006). Polarity correspondence: A general principle for performance of speeded binary classification tasks. *Psychological bulletin*, *132*(3), 416-442. https://doi.org/10.1037/0033-2909.132.3.416

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 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

Reber, R., Winkielman, P., & Schwarz, N. (1998). Effects of perceptual fluency on affective judgments. *Psychological Science*, *9*(1), 45–48. https://doi.org/10.1111/1467-9280.00008

Schwarz, N. (2004). Metacognitive experiences in consumer judgment and decision making. *Journal of Consumer Psychology*, *14*, 332-348. https://doi.org/10.1207/s15327663jcp1404_2

Singmann, H., Bolker, B., Westfall, J., Aust, F. & Ben-Shachar, M. S. (2021). afex: Analysis of Factorial Experiments. R package version 1.0-1. https://CRAN.R-project.org/package=afex

Sommet, N, & Morselli, D. (2017). Keep calm and learn multilevel logistic modeling: A simplified three-step procedure using Stata, R, Mplus, and SPSS. *International Review of Social Psychology*, *30*, 203-218. http://doi.org/10.5334/irsp.90

Stump, A., Rummel, J., & Voss, A. (2022). Is it all about the feeling? Affective and (meta-)cognitive mechanisms underlying the truth effect. *Psychological Research, 86*, 12-36. https://doi.org/10.1007/s00426-020-01459-1

Unkelbach, C. (2006). The learned interpretation of cognitive fluency. *Psychological Science*, *17*, 339–345. https://doi.org/10.1111/j.1467-9280.2006.01708.x

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, NY: Psychology Press.

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., & 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., & 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

Warriner, A.B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavioral Research*, *45*, 1191-1207. https://doi.org/10.3758/s13428-012-0314-x

Wentura, D. (2000). Dissociative affective and associative priming effects in the lexical decision task: Yes versus no responses to word targets reveal evaluative judgment tendencies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(2), 456–469. https://doi.org/10.1037/0278-7393.26.2.456

Winkielman, P., & Cacioppo, J. T. (2001). Mind at ease puts a smile on the face: Psychophysiological evidence that processing facilitation elicits positive affect. *Journal of personality and social psychology*, *81*(6), 989-1000. https://doi.org/10.1037/0022-3514.81.6.989

Winkielman, P., Schwarz, N., Fazendeiro, T. A., & Reber, R. (2003). The hedonic marking of processing fluency: Implications for evaluative judgment. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 189–217). Lawrence Erlbaum Associates Publishers.