Detecting Falsehood Relies on Mismatch Detection between Sentence Components

Rebecca Weil[a*] & Liad Mudrik[b,c]

[a] Department of Psychology, Faculty of Health Sciences, University of Hull, HU6 7RX, United Kingdom; r.weil@hull.ac.uk; [*]Corresponding author

[b] School of Psychological Sciences, Faculty of Social Sciences, Tel Aviv University, Tel Aviv 69978, Israel; mudrikli@tau.ac.il

[c] Sagol School of Neuroscience, Tel Aviv University, Tel Aviv 69978, Israel

Wordcount: [12107]

Abstract

How do people process and evaluate falsehood of sentences? Do people need to compare presented information with the correct answer to determine that a sentence is false, or do they rely on a mismatch between presented sentence components? To illustrate, when confronted with the false sentence 'trains run on highways', does one need to know that trains do not run on highways or does one need to know that trains run on tracks to reject the sentence as false? To investigate these questions, participants were asked to validate sentences that were preceded by images (Experiments 1-3) conveying a truth-congruent or a falsehood-congruent component of the sentence (e.g., an image of tracks/highway preceding the sentence 'trains run on tracks/highways') or by words (Experiment 4) that were either sentence-congruent, truth-congruent, or both (e.g., the word 'train/tracks' preceding the sentence 'trains run on tracks/highways'). Results from four experiments showed that activating sentence-congruent concepts facilitates validation for both false and true sentences but that activating truth-congruent concepts did not aid the validation of false sentences. The present findings suggest that a detection of falsehood relies on a mismatch detection between sentence's components, rather than on the activation of true content in the context of a particular sentence.

Keywords: validation, falsehood, false information, congruency, priming

1. Introduction

We live in an age of misinformation (Helfand, 2016). Fake news spread fast and wide (Lazer, et al., 2018), and seem difficult to detect (e.g., Conroy, Rubin, & Chen, 2015) and debunk (see Chan, Jones, Hall Jamieson, & Albarracín, 2017). People adopt false information even when they have existing knowledge that should have allowed them to reject it (for an overview, see Rapp & Braasch, 2014). This is even more striking given previous research on validation, showing that people are proficient at detecting falsehood (Cook & O'Brien, 2014; Isberner & Richter, 2013, 2014; Richter, Schroeder, & Wöhrmann, 2009). Arguably, the prominence of false information might be taken as failures of validation processes (e.g., Pantazi, Kissine, & Klein, 2018), whose mechanisms are not well-enough understood (Kendeou, 2014). Thus, it seems crucial to investigate how the mental system validates information to explain and prevent such failures.

The present research accordingly focuses on validation processes: when presented with a false sentence (e.g., 'macadamia are berries') do people need to compare presented information with the correct answer (e.g., 'nuts') to determine that a sentence is false, or do they rely on a mismatch between the sentence components to detect falsehood (e.g., the mismatch between 'macadamia' and 'berries')?

Contemporary models of validation in text-comprehension suggest that validation is a by-product of comprehension (cf. Gilbert, 1991; Connell & Keane, 2006) and emphasize the importance of prior knowledge to the process (Cook & O'Brien, 2014; O'Brien & Cook, 2016; Richter, 2015; Singer, 2013). The Resonance-Integration-Validation (RI-Val) model (Cook & O'Brien, 2014; O'Brien, & Cook, 2016) postulates three asynchronous processes that are part of comprehension. In a first resonance stage, incoming information leads to active representations in memory. These representations include both the newly encoded information, and passively activated long-term-memory components that are related to that information (e.g., general world knowledge). For example, when one reads the sentence

'macadamia are berries' the sentence-components (macadamia, berries) are held to be activated together with associations like 'cookies', 'Hawaii', 'nuts', 'healthy' or 'fruit'. In a second stage, all activated components are integrated, or linked to each other on the basis of general conceptual overlap or goodness of fit. For example, macadamia might be linked to berries when both concepts are recognized as food. In a third stage, formed linkages are validated against contents from long-term memory by a passive, pattern-matching process. At this stage, the mismatch between the active link 'macadamia – berries' and the link from long-term memory 'macadamia – nuts' should be detected. Accordingly, prior knowledge is essential to detect that information is false.

Similarly, Richter et al. (2009) showed that relevant background knowledge is indeed used to validate information. Participants were able to routinely reject information as false (e.g., 'Soap is edible') when they held relevant background beliefs as compared to when they did not have such knowledge (e.g., 'Toothpaste contains sulfur'). In line with these findings, validation might be described as an evaluative process that compares incoming information with existing knowledge (Richter, 2015). When a mismatch between incoming information and existing knowledge is detected, information is rejected as false. Accordingly, the accessibility of background knowledge is a precondition for successful validation (Richter et al., 2009). Thus, it seems beyond dispute that existing knowledge plays an important role for validation processes (Singer, 2019; see also Singer & Doehring, 2014). However, this does not necessarily imply that background knowledge needs to consist of a correct answer or corresponding true concepts (in the context of the present paper, by 'true' we mean 'in accord with general knowledge and well-known facts' rather than in accord with reality, in line with coherence theories of truth; for an overview see Kirkham, 1992). Some information (e.g. 'Soap is edible') does not have a corresponding true concept to compare it to and hence, validation might entail having knowledge about semantic network affiliation (e.g., Soap belongs to the category of hygiene products but not to the category of food). Accordingly, the

sentence components ‘soap’ and ‘edible’ create a mismatch, that once detected leads to the conclusion that the sentence is false. Other information (‘macadamia are berries’) might have a corresponding true concept (‘nuts’) but could also be validated according to semantic network affiliation (i.e., one knows that macadamia are nuts *and* that they do not belong to the category of berries). Accordingly, falsehood could be discovered due to the mismatch detection of the components ‘macadamia’ and ‘berries’ or due to the mismatch detection of the true knowledge (‘nuts’) and the presented information (‘berries’).

When different methods of validation (i.e., mismatch detection between sentence components or between true concept and presented information) can be utilized, comprehenders might engage in minimal semantic processing that is just good enough to complete the validation process. In line with the Good-Enough Representations approach (Ferreira, Bailey, & Ferraro, 2002) comprehenders do not always engage in complete and detailed processing of a sentence, so that the latter only occurs if it is required. Thus, if the detection of a conceptual mismatch between sentence components is enough to determine the validity of a sentence, comprehenders might not access true concepts even if they are available (for a similar discussion see Richter & Maier, 2017). More support for this claim comes from the Discrepancy-Induced Source Comprehension (D-ISC) model (Braasch & Bråten, 2017), suggesting that the detection of a conceptual mismatch or conflict motivates comprehenders to invest mental effort and strategically use background knowledge to resolve the conflict. When the main goal is to determine the validity of a sentence, comprehenders might decide not to spend additional effort as the detection of a conflict is enough to determine that a sentence is false. This reasoning is in accordance with semantic integration studies (e.g., Berkum, Hagoort, & Brown, 1999), which point at the sensitivity of the mental system to semantic violations (e.g., Kutas & Hillyard, 1980) and incongruencies (e.g., Hagoort, Hald, Bastiaansen, & Petersson, 2004; see also Biderman & Mudrik, 2017, for implicit processing of incongruencies), typically indexed by the N400 component (for review,

see Kutas & Federmeier, 2011). Critically, these studies demonstrated the importance of association strength in processing, irrespective of validity (see also DeLong, Urbach & Kutas, 2005). For example, the sentences 'cows drink milk' and 'cows drink juice' are both false. If validation always involves comparing false concepts (i.e. 'milk', 'juice') against the true concept (i.e. 'water') both sentences should be easily detected as false. However, the sentence 'cows drink juice' elicits a greater N400 effect than the sentence 'cows drink milk' (see Kutas & Hillyard, 1984) and may be more difficult to explicitly identify as false (see Hinze, Slaten, Horton, Jenkins, & Rapp, 2014). This might be the case because 'milk' and 'cow' are part of the same semantic associative network but 'juice' is not (see also Erickson, & Mattson, 1981; Sanford, 2002). Thus, 'cows drink juice' might be validated according to a lack of semantic overlap between sentence components, while the validation of 'cows drink milk' would require a different strategy (see also Cook, Walsh, Bills, Kircher, & O'Brien, 2016).

In this study, we manipulated the likelihood of using different strategies (i.e., mismatch detection between sentence components or between true concept and presented information) to assess the validity of a sentence that could either be true (e.g., 'macadamia are nuts') or false (e.g., 'macadamia are berries'). We used priming to pre-activate concepts that are either congruent with true concepts (e.g., 'nuts'), with false concepts (e.g., 'berries') or with a component of the sentence (e.g., 'macadamia'). Both speed and accuracy of validation (i.e., determining if a sentence is true of false) were measured, following the different primes. Firstly, we investigated if pre-activating sentence-related concepts (i.e., true concepts or sentence components) by means of priming can facilitate explicit validity judgments of true and false sentences; going beyond previous studies which either (a) did not pre-activate knowledge but relied on participants' existing knowledge in long-term memory (e.g., Richter et al., 2009) when investigating validation processes, or (b) examined the effects of pre-activating concepts on stimuli processing irrespective of validation attempts (e.g., Kutas & Hillyard, 1984; see again Kutas & Federmeier, 2011 for review). Secondly, we investigated

the unique contribution of detecting the mismatch between sentence components vs. detecting the mismatch between the true concept and sentence components to validation processes.

For true sentences, pre-activation of any sentence-congruent concept should facilitate validation. Here, truth-congruent concepts are both congruent with sentence content as well as with related knowledge from long-term memory, while falsehood-congruent concepts are neither and thus, should not facilitate the validation process for true sentences. A more interesting case is posed by false sentences, where truth-congruent concepts are incongruent with the sentence content, but activate related knowledge (e.g., 'nuts' in the above example is a concept that represents the true information that does not appear in the sentence).

Notably, activating both truth-congruent concepts and falsehood-congruent concepts prior to processing a false sentence might facilitate validation processes. For falsehood-congruent concepts, activation of sentence-congruent information might (a) facilitate sentence comprehension, by pre-activating one of the components of the sentence, and, consequently, (b) facilitate detection of semantic network affiliation between the sentence components. Much like decisions about a target (e.g., a word/non-word decision; Meyer & Schvaneveldt, 1971) that are facilitated when the target is preceded by the same or a semantically related prime (Neely, 1977; Posner & Snyder, 1975), activating a semantic association that is congruent with the false concept might lead to a faster mismatch detection between sentence components (e.g., that macadamia do not belong to the category of berries).

Activating truth-congruent concepts might also facilitate validation processes. Arguably, truth-congruent concepts allow for an immediate comparison between the content of the false sentence and the true background knowledge (O'Brien, & Cook, 2016; Richter 2015). We accordingly reasoned that several patterns of results might be found, each implying a different theoretical account of the role of true concepts in validation processes. If truth-congruent but not falsehood-congruent primes facilitate validation processes of false sentences, access to true concepts from long-term memory might be necessary to validate

false information. This would imply that activating sentence-congruent information is not enough to affect performance in a validation task. Alternatively, if falsehood-congruent but not truth-congruent concepts speed up validation, it would imply that knowledge about semantic network affiliation suffices to detect falsehood. That is, what is needed for validation is the ability to detect a mismatch between the components of the false sentence, rather than the activation of the true information the sentence does not portray. And so, primes that already activate one of the sentence's components facilitate its comprehension and the comparison between semantic affiliations. Such an outcome would also imply that truth-congruent primes, although associatively related, might not activate sentence components strongly enough to facilitate validation. Yet another possibility is that the pre-activation of *both* falsehood- and truth-congruent concepts facilitates validation. In such a case, differences between facilitation- strengths would indicate whether validation can be carried out on the basis of both access to true concepts from long-term memory and knowledge about semantic network affiliation (i.e., no difference between truth-congruent and falsehood-congruent primes) or whether truth-congruent primes contribute to detection of semantic network affiliation (i.e., falsehood-congruent primes lead to stronger facilitation than truth-congruent primes). A third alternative is that truth-congruent primes lead to stronger facilitation effects compared to falsehood-congruent primes. Notably, truth-congruent concepts are part of a respective semantic network, and so it could be argued that pre-activation of any part of the semantic network should be helpful to determine semantic network affiliation. Still, the pre-activation of truth-congruent concepts might aid a unique mechanism underlying validation processes, over and above the pre-activation of concepts in the semantic network. As outlined earlier, this unique mechanism might be necessary for the discovery of falsehood: arguably, detection of the mismatch between information presented in a sentence and true concepts from long-term memory might be needed to determine that the sentence is false. If this is the

case, the pre-activation of truth-congruent concepts should aid validation, over and above pre-activation of any other concepts in the semantic network.

All of the above outlined results are compatible with existing models of validation (e.g., Cook & O'Brien, 2014; O'Brien, & Cook, 2016; Richter et al. 2009; Richter, 2015), which emphasize the importance of background information in validation processes. Importantly however, each result will have a different implication regarding the contribution of detecting the mismatch between sentence components vs. between the true concept and presented information. Accordingly, the goals of this research are twofold: a) to assess the effect of knowledge pre-activation on validation processes b) to examine if different types of knowledge exert a different effect on these processes.

2. Experiment 1

To investigate whether pre-activating truth-congruent/falsehood-congruent concepts facilitates validation processes for false sentences, participants were presented with sentences, and asked to judge their validity (i.e., determine whether they are true or false). We assumed that successful validation should translate into 'false' judgments when the sentence is false and into a 'true' judgment when the sentence was considered true (see Isberner & Richter, 2014; Richter, 2015; Richter et al., 2009). Activation of truth-congruent or falsehood-congruent concepts in the context of these sentences was manipulated using picture primes (see Orenes & Santamaría, 2014), depicting concepts that were either truth-congruent, falsehood-congruent or unrelated to the sentences' content. For example, the sentence 'trains run on tracks' or 'trains run on highways' could have been preceded by an image of tracks (truth-congruent), a highway (falsehood-congruent) or a TV test pattern (unrelated). If activating background knowledge can facilitate explicit validation processes, either truth-congruent primes or falsehood-congruent primes (or both) should facilitate validation for false sentences (i.e., helps participants determine that a sentence is false). Thus, falsehood

judgments should be faster and more accurate when preceded by primes depicting either truth-congruent or falsehood-congruent concepts as compared to unrelated primes. In case such a facilitation is found, one could then inspect its magnitude for the two types of primes.

Notably, it might be questioned whether priming can selectively activate concepts that are either congruent with true or with false concepts, but not with both. Consider, for example, the sentence 'Fire is cold'. Pre-activating the true concept 'hot' will activate the associatively related concept 'fire' to some degree, but also the semantically related antonym 'cold' (for an overview see McNamara, 2005). Nevertheless, true concepts should be activated more strongly by truth-congruent primes as compared to falsehood-congruent primes, and false concepts should be activated more strongly by falsehood-congruent primes as compared to truth-congruent primes (e.g., Hutchison, 2003; Traxler, Foss, Seely, Kaup, & Morris, 2000). Thus, the two different prime types have the potential to produce meaningful differences with respect to the facilitation of validation processes.

2.1. Methods

2.1.1. Participants and design. Eighty undergraduates at the University of Hull (52 female, 27 male, 1 not reported; M_{age} = 19.65; SD_{age} = 2.05) participated in an online study on 'judgments and visual distraction' in return for course credit. [1] Due to the nature of the experiment being conducted online, in a post-experimental demographic questionnaire, participants were asked whether they were native English speakers, interrupted during the experiment or in the presence of others while performing the task, and whether they had any educated guess concerning the purpose of the experiment.

[1] The experiments were approved by the ethics committee of the University of Hull, and informed consent was obtained before participants started the task. As this is a new paradigm, we determined the sample size for each study beforehand, with the requirement of at least 80 participants in Experiment 1 and 3, based on the availability of participants in the department's subject pool. We aimed to recruit as many participants as were available during the term of the study. The sample sizes of the online Experiments 2 and 4 were set to 100, accounting for potentially incomplete submissions, due to the length of the study. Sensitivity analyses (GPower 3.1.9.2), assuming a power of (1-β) = .80, revealed that the experiments were sensitive to detect effect sizes of $\eta_p^2 > .02$, for the main statistical effects of interest. We collected the data for each experiment in one shot without prior statistical analyses. We report all data exclusions, all manipulations, and all measures. Materials and data are available at https://osf.io/c6j4b/.

The study consisted of a 2 (Sentence Validity: true vs. false) × 3 (Prime Congruency: truth-congruent vs. falsehood-congruent vs. unrelated) within-participants design. Stimulus presentation and response collection were controlled by Inquisit 5.0.11.0.

2.1.2. Stimuli. We created easy sentences, involving simple declarative facts, with universally correct answers (for an overview of experimental sentences see Appendix A). In total, 480 sentences were created, half of them true (e.g., 'Cheetahs run fast') the other half false (e.g., 'Turtles move fast'). Each participant saw 120 different true and 120 different false sentences out of the total number of sentences. A participant never saw both the true (e.g., 'Cheetahs run fast') and the false pairing (e.g., 'Cheetahs run slow') of the same sentence with a concept. Whether a sentence appeared with a true or false concept was counterbalanced between participants. All sentences had the same general structure, whereby the concept was presented at the end of the sentence. The mean number of words per sentence was 5.15 (*SD* = 1.86). All sentences were presented in random order.

Primes were pictures of objects or events, taken from Internet resources, that signified concepts of the upcoming sentences or were unrelated to them (i.e., a TV test pattern). Pictures were selected to signify the truth-congruent and falsehood-congruent concepts in sentences as literally as possible. We ensured that pictures did not show the subject of the sentence (or elements of it) whenever this was possible. For example, the picture that represented the truth-congruent concept for the sentence 'Broccoli is a vegetable' showed several vegetables (e.g., carrots, bell pepper, kale) not including broccoli. When a literal representation was not possible (e.g., for the concept 'cheap') close metaphorical representations were chosen (e.g., a price tag showing a % sign).

In total, 240 different pictures were created, corresponding to the concepts mentioned in the 240 true sentences, 76 out of 240 were metaphorical representations (see Appendix A). For example, a picture of a stopwatch was created to match the true sentence 'Cheetahs run fast' or the false sentence 'Turtles move fast'. Primes and sentences were matched in a way

that one third of sentences (80 pictures) were preceded by congruent primes, one third (80 pictures) was preceded by incongruent primes and one third (80 pictures) by unrelated primes. For true sentences (e.g., 'Cheetahs run fast'), congruent primes (e.g., stopwatch; 40 pictures) matched the true concept mentioned in the sentence (e.g., fast) and incongruent primes (e.g., hourglass; 40 pictures) represented a potentially false concept (e.g., slow). For false sentences (e.g., 'Turtles move fast'), congruent primes (e.g., stopwatch; 40 pictures) matched the false concept mentioned in the sentence (e.g., fast) and incongruent primes (e.g., hourglass; 40 pictures) represented a true concept (e.g., slow). The assignment of congruent, incongruent and unrelated primes to sentences was fully counterbalanced between participants.

2.1.3. Procedure. Participants saw 240 trials. In each trial, they were asked to judge a different sentence as either true or false. Participants were informed that coherent sentences about unreal or fictional events (e.g., 'Dragons breathe fire') should be considered as true. They were instructed that first, a picture would briefly appear on the screen and that they should do nothing in response to the picture, as it signaled that the sentence is about to appear. The picture referred to concepts that were either truth-congruent, falsehood-congruent or unrelated to the sentences' content.

Each trial started with a warning signal (+++), presented in the center of the screen for 500 ms. Subsequently, the picture was presented for 500 ms, followed by a true or false sentence, separated by a 50 ms blank screen. Participants were instructed to indicate whether the sentence was false, by pressing 'q', or true, by pressing 'p', on their keyboard. Labels were shown at the bottom of the screen to remind participants about the key assignment. The sentence stayed on the screen until participants indicated their answer (see Figure 1). Each trial was separated by a 1000 ms interval.

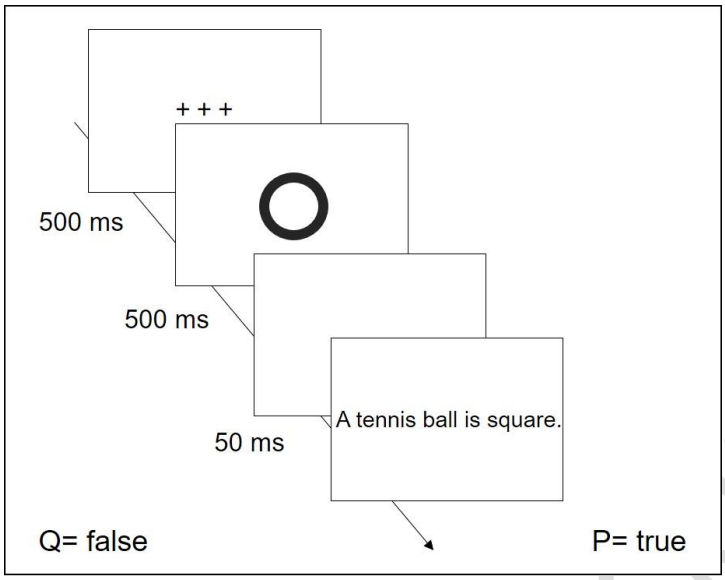

Figure 1: Example of trial sequence in Experiment 1, 2 and 3 (note that in Experiments 2 and 3, the structure of the sentence was different; applied to this example, it would be 'Square is the shape of a tennis ball'.)

2.2. Results

12.5% of participants reported being non-native English speakers, 3.8% were interrupted during the experiment and 33.8% were in the presence of others while performing the task. These participants are included in the following analyses. Yet, to investigate whether results are affected, we excluded these participants in a separate analysis. Results show that the general pattern of effects stays the same.

We only included trials with reaction times above 300 ms and below 10000 ms (97.92% of trials), assuming that correct judgements faster than 300 ms are driven by anticipations rather than reflecting validation, and responses slower than 10000 ms might indicate that participants were not concentrated on the task in the respective trial. We chose a relatively high upper threshold for truncation to account for differences in sentence length and reading speed. Data from one participant were incompletely recorded and were not included in the analyses. The following analyses are accordingly based on 79 participants.

2.2.1. Reaction times. We excluded all trials in which sentences were judged incorrectly (9.58% of trials). To test whether pre-activating truth-congruent or falsehood-congruent concepts influenced the validation latency, we conducted linear mixed-model analyses (Baayen, Davidson, & Bates, 2008; see also Clark, 1973) using the lme4 package (Bates, Mächler, Bolker, & Walker, 2015) in the statistical software R Version 3.5.1 for Windows (R Core Team, 2018). We report in the following models with the maximal random-effect structure that converged (Barr, Levy, Scheepers, & Tily, 2013). For all models, we used reference-coding with unrelated primes as reference level for Prime Congruency. That is, we compared each level of Prime Congruency to the reference level. The intercept represents the cell mean of the unrelated primes.

We fitted a model including Sentence Validity, Prime Congruency and the two-way interaction of Sentence Validity and Prime Congruency as fixed effects. The model featured by-subject and by-item random intercepts, as well as a by-subject and by-item random slopes for Sentence Validity. The analysis revealed a main effect of Sentence Validity, $\chi^2(1) = 9.02$, $p = .003$, a main effect of Prime Congruency, $\chi^2(2) = 19.57$, $p < .001$, and a two-way interaction of Sentence Validity and Prime Congruency, $\chi^2(2) = 36.84$, $p < .001$ (see Figure 2, left). Simple contrasts were calculated using the R emmeans package based on the R lsmeans package (Lenth, 2016). The simple contrast showed that, for true sentences, participants judged sentences faster when they were preceded by truth-congruent primes ($M_{EM} = 1762.08$, $SE = 73.38$) as compared to unrelated primes ($M_{EM} = 1851.30$, $SE = 73.38$) and judged sentences slower when they were preceded by falsehood-congruent primes ($M_{EM} = 1920.32$, $SE = 73.51$; see Table 1), compared to unrelated primes. For false sentences, participants judged sentences faster when they were preceded by falsehood-congruent primes ($M_{EM} = 1874.02$, $SE = 69.48$) as compared to unrelated primes ($M_{EM} = 1957.02$, $SE = 69.52$). There was no difference between truth-congruent primes ($M_{EM} = 1909.70$, $SE = 69.51$; see Table 1)

and unrelated primes. Thus, truth-congruent primes facilitated judgments for true sentences, while falsehood-congruent primes had different effects for true and false sentences: they slowed down judgments for true sentences but facilitated judgments for false sentences.

Table 1: Comparisons of estimated marginal means of participants' reaction times, p-values are adjusted by Tukey-method, Experiment 1.

Contrast	*Estimate*	*Standard Error*	*p-value*
True Sentences			
Unrelated vs. truth-congruent prime	89.22	23.62	< .001
Unrelated vs. falsehood-congruent prime	-69.02	24.00	.01
False Sentences			
Unrelated vs. truth-congruent prime	47.32	23.52	.11
Unrelated vs. falsehood-congruent prime	82.99	23.42	.001

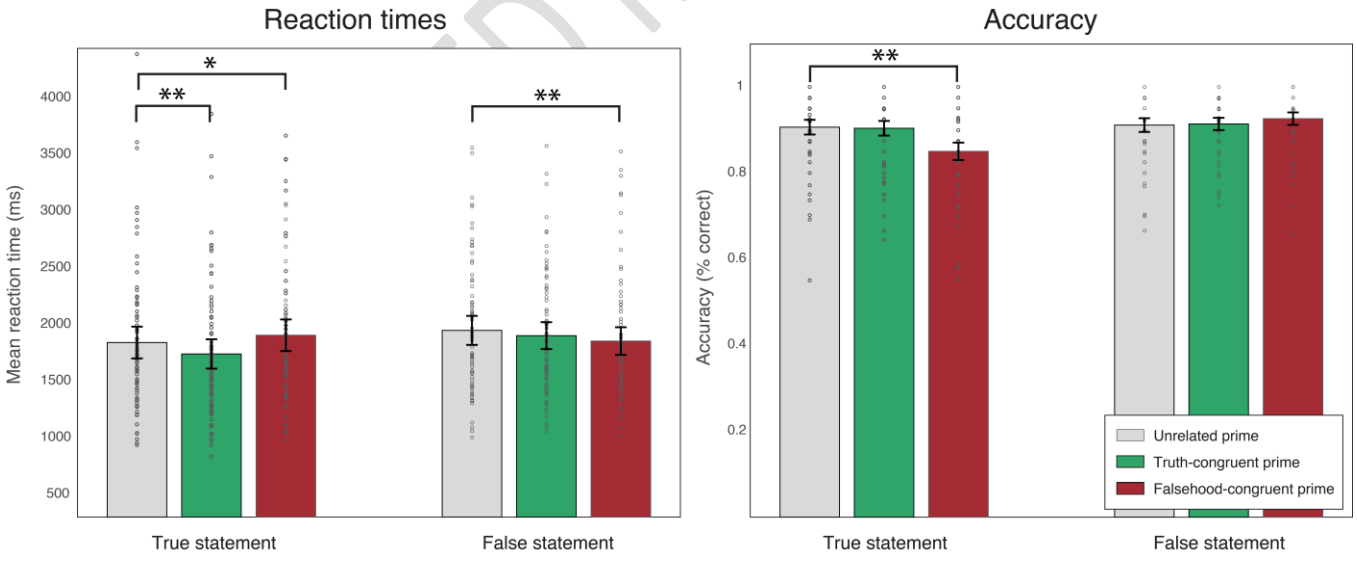

Figure 2: Mean reaction times (left) and accuracy (right) of true/false judgment as a function of Sentence Validity and Prime Congruency in Experiment 1. Error bars depict 95% confidence intervals. Circles denote data points of individual subjects. Significant contrasts are indicated (** equals $p \leq .001$; * equals $p < .05$).

2.2.2. Accuracy. To test whether pre-activating truth- or falsehood-congruent concepts influenced validation accuracy, we fitted a model including Sentence Validity, Prime Congruency and the two-way interaction of Sentence Validity and Prime Congruency as fixed effects with by-subject and by-item random intercepts. The analysis showed a main effect of Sentence Validity, $\chi^2(1) = 49.12$, $p < .001$, a main effect of Prime Congruency, $\chi^2(2) = 25.78$, $p < .001$, and a two-way interaction of Sentence Validity and Prime Congruency, $\chi^2(2) = 51.91$, $p < .001$ (see Figure 2, right). The simple contrast showed that, for true sentences, participants were less accurate following falsehood-congruent primes ($M_{prob} = .89$, $SE = .01$), as compared to unrelated primes ($M_{prob} = .94$, $SE = .01$). There was no difference in accuracy between unrelated and truth-congruent primes ($M_{prob} = .93$, $SE = .01$; see Table 2). For false sentences, accuracy of judgments did not differ between unrelated ($M_{prob} = .94$, $SE = .01$), falsehood-congruent ($M_{prob} = .95$, $SE = .01$) or truth-congruent primes ($M_{prob} = .94$, $SE = .01$; see Table 2). Thus, falsehood-congruent primes reduced correct judgments for true sentences, but not for false sentences.

Table 2: Comparisons of log odds ratio of participants' accuracy, p-values are adjusted by Tukey-method, Experiment 1.

Contrast	*Odds Ratio*	*Standard Error*	*p-value*
True Sentences			
Unrelated vs. truth-congruent prime	1.05	.10	.87
Unrelated vs. falsehood-congruent prime	1.90	.16	< .001
False Sentences			
Unrelated vs. truth-congruent prime	.98	.09	.96
Unrelated vs. falsehood-congruent prime	.82	.08	.10

2.2.3. Bayes Factor Analysis. To quantify the evidence for the presence or absence of effects we also calculated Bayes factors (BF; see Table 3), using JASP (2018). We adopted

the convention that $BF_{10} = 1$ implies no evidence for an effect (i.e., the data are as likely to occur under H0 as under H1), $1 < BF_{10} \leq 3$ implies anecdotal evidence for H1, $3 < BF_{10} \leq 10$ implies moderate evidence for H1, $10 < BF_{10} \leq 30$ implies strong evidence for H1, $30 < BF_{10} \leq 100$ implies very strong evidence for H1 and $BF_{10} > 100$ implies decisive evidence for H1(Jeffreys, 1961; Lee & Wagenmakers, 2013). Similarly, $.30 < BF_{10} \leq 1$ implies anecdotal evidence for H0, $.10 < BF_{10} \leq .30$ implies moderate evidence for H0, $.03 < BF_{10} \leq .10$ implies strong evidence for H0, $.01 < BF_{10} \leq .03$ implies very strong evidence for H0 and $BF_{10} < .01$ implies decisive evidence for H0.

Table 3: Bayes Factors for main effects, interactions and simple comparisons of participants' reaction times and accuracy, Experiment 1.

	Bayes Factors	
Effects and Simple Comparisons	*Reaction Times*	*Accuracy*
Sentence Validity	$BF_{10} = 40.88$	$BF_{10} = 56692.93$
Prime Congruency	$BF_{10} = 2.05$	$BF_{10} = 4.06$
Sentences Validity × Prime Congruency	$BF_{10} = 520.71$	$BF_{10} = 547675.89$
True sentences		
Unrelated vs. truth-congruent prime	$BF_{10} = 4.38$	$BF_{10} = .17$
Unrelated vs. falsehood-congruent prime	$BF_{10} = .62$	$BF_{10} = 3148.05$
False sentences		
Unrelated vs. truth-congruent prime	$BF_{10} = .53$	$BF_{10} = .18$
Unrelated vs. falsehood-congruent prime	$BF_{10} = 11.27$	$BF_{10} = 1.31$

2.3. Discussion

The results of Experiment 1 showed that pre-activating participants' background knowledge facilitates explicit validation: falsehood-congruent primes facilitated validity-judgments of false sentences and truth-congruent primes facilitated validity-judgements of

true sentences. This result is expected, as it is in line with validation models (e.g., Cook & O'Brien, 2014; Richter et al. 2009). Thus, the critical question here is whether activating truth-congruent concepts will also facilitate validation processes of false sentences, a possibility that can be derived from extant models of validation (e.g., O'Brien & Cook, 2016; Richter, 2015; Singer, 2019), but has never been directly tested. Here, the results suggest it does not: truth-congruent content did not facilitate validation, demonstrating the importance of knowledge about semantic network affiliations rather than knowledge about the true concept.

Yet, the above conclusion might be mitigated by an additional factor that might have influenced participants' reaction times in Experiment 1. In the current design, primes and beginning of sentences either created a match (e.g., tracks/trains) or a mismatch (highways/trains). Such matches or mismatches occurred in equal proportions for true and false sentences, and thus, did not allow participants do use these (mis)matches as a strategy to determine the validity of a sentence. Nevertheless, they might have triggered early response tendencies that could have affected performance, due to a conflict between the match/mismatch between the prime and the first word, and the validity of the sentence. To illustrate, when the sentence 'Trains run on tracks' is preceded by the prime 'highway', a mismatch is created between 'highway' and 'trains', although the actual sentence is true. Similarly, when the sentence 'Trains run on highways' is preceded by the prime 'tracks', a conflict might arise between detecting the match between 'tracks' and 'trains', and the overall falsehood of the sentence. Thus, it could be that both these conditions evoke slower reaction times than trials where the relations between the prime and the first word accord with the overall validity of the sentence. And so, one could argue that truth-congruent primes (here, 'tracks') actually facilitate validation of false sentences, but this effect is masked due to the opposite effect evoked by the conflict between the falsehood of the sentence and the congruency between the prime and the first word.

Hence, the paradigm in Experiment 1 might not have been a fair test for the role of true concepts in validation processes. To account for early triggered response tendencies and to investigate whether the effects found in Experiment 1 are robust, we changed the structure of sentences in Experiment 2. In the new structure, truth-congruent primes would evoke a mismatch with the first word of a false sentence, so no conflict should emerge; hence, if truth-congruent primes indeed facilitate validation processes, there will be no opposite effect that would counteract this facilitation.

3. Experiment 2

To test the reproducibility and robustness of our findings, the words in the sentences were now switched, so that concepts that ended the sentences in Experiment 1 were presented at the beginning of the sentences (e.g., 'tracks are the infrastructure trains run on' for a true sentence, and 'highways are the infrastructure trains run on' for a false one; see Appendix B). Thus, for false sentences, falsehood-congruent primes matched now with the beginning of the false sentences (e.g., a picture of a highway and the word 'highway') and truth-congruent primes mismatched with the beginning of false sentences (e.g., a picture of tracks and the word 'highway'). Accordingly, and differently from Experiment 1, a potential facilitation of truth-congruent concepts for false sentences does not require to overcome a reverse response tendency.

3.1. Methods

3.1.1. Participants and design. We recruited 100 participants (55 female, 41 male, 2 other, 2 not reported; M_{age} = 35.30; SD_{age} = 12.84) via Prolific Academic (see Palan, & Schitter, 2018; Peer, Brandimarte, Samat, & Acquisti, 2017). Participants could only sign up for the experiment if they resided in the United States and were native English speakers. We required that they had previously completed at least 50 tests via Prolific Academic and held a record of supplying acceptable data at least 95% of the time. They received £2.08 (approx.

$2.50) for their participation. The design and post-experimental questionnaire were identical to Experiment 1.

3.1.2. Procedure and Stimuli. The procedure was identical to Experiment 1. However, true and false sentences had a different structure as compared to the first experiment. For all experimental sentences, the concept that initially ended the sentences, was now mentioned at the beginning of each sentence (see Appendix B). The mean number of words per sentence was 7.00 (*SD* = 1.58). All sentences had the same general structure, mentioning the concept before the object. This was the case even if the concept was not the first word of the sentence (e.g., 'an example for a fruit is a pear'; sentence in Experiment 1: 'a pear is a fruit'). The mean number of words between prime and concept was .88 (*SD* = 1.36). The overall meaning of the sentences was not changed.

3.2. Results

2% of participants reported being non-native English speakers, 1% were interrupted during the experiment and 2% were in the presence of others while performing the task. These participants are included in the following analyses. To investigate whether results are affected, we excluded these participants in a separate analysis. Results show that the general pattern of effects stays the same. We only included trials with reaction times above 300 ms and below 10000 ms (98.28% of trials). Data from two participants were not recorded. The following analyses are based on 98 participants.

3.2.1. Reaction times. We excluded all trials in which sentences were judged incorrectly (7.69% of trials). To test whether pre-activating truth- or falsehood-congruent concepts influenced reaction times of validity-judgments when the concept is mentioned at the beginning of the sentence, we fitted a model including Sentence Validity, Prime Congruency and the two-way interaction of Sentence Validity and Prime Congruency as fixed effects. The model featured by-subject and by-item random intercepts, as well as a by-subject and by-item random slopes for Sentence Validity. The analysis showed a main effect for Sentence

Validity, $\chi^2(1) = 10.97$, $p < .001$, a main effect for Prime Congruency, $\chi^2(2) = 24.57$, $p < .001$, and a two-way interaction of Sentence Validity and Prime Congruency, $\chi^2(2) = 80.20$, $p < .001$ (see Figure 3, left). Simple contrast showed that, for true sentences, the same patterns of results as in Experiment 1 was found, so that participants judged sentences faster when they were preceded by truth-congruent primes (M_{EM} = 1918.24, SE = 63.68) as compared to unrelated primes (M_{EM} = 2021.43, SE = 63.70) and judged sentences slower when they were preceded by falsehood-congruent primes (M_{EM} = 2115.27, SE = 63.75; see Table 4). For false sentences, as in Experiment 1, only falsehood-congruent primes seemed to affect performance: participants judged sentences faster when they were preceded by falsehood-congruent primes (M_{EM} = 2048.21, SE = 67.83) as compared to unrelated primes (M_{EM} = 2151.94, SE = 67.81), and there was no difference between truth-congruent primes (M_{EM} = 2118.09, SE = 67.83) and unrelated primes (see Table 4). Thus, when the content of the prime matched the content of the sentence, validity-judgments were faster, irrespective of the sentences' validity (see also Table 6).

Table 4: Comparisons of estimated marginal means of participants' reaction times, p-values are adjusted by Tukey-method, Experiment 2.

Contrast	*Estimate*	*Standard Error*	*p-value*
True Sentences			
Unrelated vs. truth-congruent prime	103.19	21.71	< .001
Unrelated vs. falsehood-congruent prime	-93.85	21.96	< .001
False Sentences			
Unrelated vs. truth-congruent prime	33.85	21.67	.26
Unrelated vs. falsehood-congruent prime	103.73	21.68	< .001

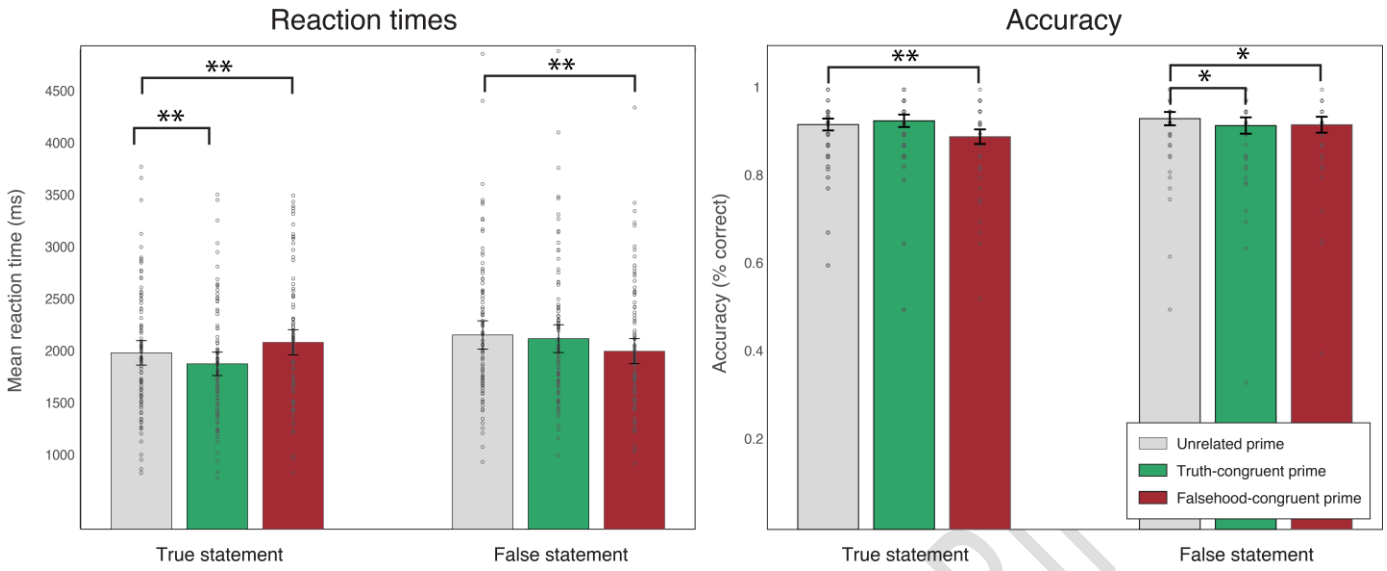

Figure 3: Mean reaction times (left) and accuracy (right) of true/false judgment as a function of Sentence Validity and Prime Congruency in Experiment 2. Error bars depict 95% confidence intervals. Circles denote data points of individual subjects. Significant contrasts are indicated (** equals $p \leq .001$; * equals $p < .05$).

3.2.2. Accuracy. We fitted a model, including Sentence Validity, Prime Congruency and the two-way interaction of Sentence Validity and Prime Congruency as fixed effects with by-subject and by-item random intercepts. The analysis revealed a main effect of Sentence Validity, $\chi^2(1) = 13.90$, $p < .001$, a main effect of Prime Congruency, $\chi^2(2) = 35.34$, $p < .001$, and a two-way interaction of Sentence Validity and Prime Congruency, $\chi^2(2) = 21.35$, $p < .001$ (see Figure 3, right). The simple contrast showed that, for true sentences, participants were less accurate following falsehood-congruent primes ($M_{prob} = .93$, $SE = .01$), as compared to unrelated primes ($M_{prob} = .95$, $SE = .01$). There was no difference in accuracy between unrelated and truth-congruent primes ($M_{prob} = .96$, $SE = .004$; see Table 5). For false sentences, accuracy of judgments for both truth-congruent ($M_{prob} = .95$, $SE = .01$) and falsehood-congruent primes ($M_{prob} = .95$, $SE = .01$) was actually lower than for unrelated primes ($M_{prob} = .96$, $SE = .004$; see Table 5). Thus, falsehood-congruent primes reduced

accuracy for both true and false sentences and truth-congruent primes reduced accuracy for false sentences (see also Table 6).

Table 5: Comparisons of log odds ratio of participants' accuracy, p-values are adjusted by Tukey-method, Experiment 2.

Contrast	*Odds Ratio*	*Standard Error*	*p-value*
True Sentences			
Unrelated vs. truth-congruent prime	.86	.08	.20
Unrelated vs. falsehood-congruent prime	1.52	.13	< .001
False Sentences			
Unrelated vs. truth-congruent prime	1.28	.12	.02
Unrelated vs. falsehood-congruent prime	1.27	.12	.03

Table 6: Bayes Factors for main effects, interactions and simple comparisons of participants' reaction times and accuracy, Experiment 2.

	Bayes Factors	
Effects and Simple Comparisons	*Reaction Times*	*Accuracy*
Sentence Validity	BF_{10} = 3015.65	BF_{10} = .75
Prime Congruency	BF_{10} = .33	BF_{10} = 18.54
Sentences Validity × Prime Congruency	BF_{10} = 1.511e+6	BF_{10} = 6.15
True sentences		
Unrelated vs. truth-congruent prime	BF_{10} = 129.70	BF_{10} = .54
Unrelated vs. falsehood-congruent prime	BF_{10} = 36.97	BF_{10} = 13.26
False sentences		
Unrelated vs. truth-congruent prime	BF_{10} = .38	BF_{10} = 3.53
Unrelated vs. falsehood-congruent prime	BF_{10} = 10.51	BF_{10} = 2.99

3.3. Discussion

Experiment 2 replicated the results of Experiment 1 for reaction times, demonstrating their reproducibility and robustness. Truth-congruent primes facilitated judgments for true sentences, while falsehood-congruent primes slowed reactions down. Similarly, falsehood-congruent primes facilitated judgments for false sentences, providing more evidence for the role of semantic network affiliations for validation processes. However, the pattern of errors, showing that both falsehood-congruent primes and truth-congruent primes reduced accuracy for false sentences, is hard to explain. Note that the Bayes factor analysis indicated only anecdotal to moderate evidence for reduced accuracy in both cases (BF_{10} = 3.53 for truth-congruent primes; BF_{10} = 2.99 for falsehood-congruent primes). Because the error pattern does not fit with neither semantic integration accounts (e.g., Berkum, Hagoort, & Brown, 1999), nor with models of validation (e.g., O'Brien, & Cook, 2016; Richter et al. 2009), Experiment 3 aimed to replicate the results of Experiment 2 to investigate whether found effects are genuine or possible false-positives.

4. Experiment 3

To investigate the replicability of the result pattern in Experiment 2 we ran an exact replication of Experiment 2, with a new sample of participants, the same materials, procedure, and experimental design.

4.1. Methods

4.1.1. Participants and design. One-hundred-and-five undergraduates at the University of Hull (81 female, 24 male; M_{age} = 20.94; SD_{age} = 5.23) participated in an online study on 'judgments and visual distraction' in return for course credit. We limited participation to participants who had not participated in Experiment 1. The design, procedure and all materials were identical to Experiment 2.

4.2. Results

5.7% of participants reported being non-native English speakers, 3.8% were interrupted during the experiment and 26.7% were in the presence of others while performing the task. These participants are included in the following analyses. To investigate whether results are affected, we excluded these participants in a separate analysis. Results show that the general pattern of effects stays the same.

We only included trials with reaction times above 300 ms and below 10000 ms (98.99% of trials). The following analyses are based on 105 participants.

4.2.1. Reaction times. We excluded all trials in which sentences were judged incorrectly (10.30% of trials). We fitted a model including Sentence Validity, Prime Congruency and the two-way interaction of Sentence Validity and Prime Congruency as fixed effects. The model featured by-subject and by-item random intercepts, as well as a by-subject and by-item random slopes for Sentence Validity. The analysis showed a main effect of Sentence Validity, $\chi^2(1) = 12.69$, $p < .001$, a main effect of Prime Congruency, $\chi^2(2) = 30.26$, $p < .001$, and a two-way interaction of Sentence Validity and Prime Congruency, $\chi^2(2) = 111.92$, $p < .001$ (see Figure 4, left; see also Table 9). Simple contrast showed that, for true sentences, participants judged sentences faster when they were preceded by truth-congruent primes (M_{EM} = 1892.48, SE = 54.12) as compared to unrelated primes (M_{EM} = 1999.20, SE = 54.15) and judged sentences slower when they were preceded by falsehood-congruent primes (M_{EM} = 2080.24, SE = 54.27; see Table 7). For false sentences, participants judged sentences faster when they were preceded by falsehood-congruent primes (M_{EM} = 2004.11, SE = 56.53) as compared to unrelated primes (M_{EM} = 2130.60, SE = 56.54). There was no difference between truth-congruent primes (M_{EM} = 2092.53, SE = 56.57; see Table 7) and unrelated primes, replicating the results from Experiment 1 and Experiment 2.

Table 7: Comparisons of estimated marginal means of participants' reaction times, p-values are adjusted by Tukey-method, Experiment 3.

Contrast	*Estimate*	*Standard Error*	*p-value*
True Sentences			
Unrelated vs. truth-congruent prime	106.72	19.01	< .001
Unrelated vs. falsehood-congruent prime	-81.04	19.43	< .001
False Sentences			
Unrelated vs. truth-congruent prime	38.04	18.98	.11
Unrelated vs. falsehood-congruent prime	126.49	18.86	< .001

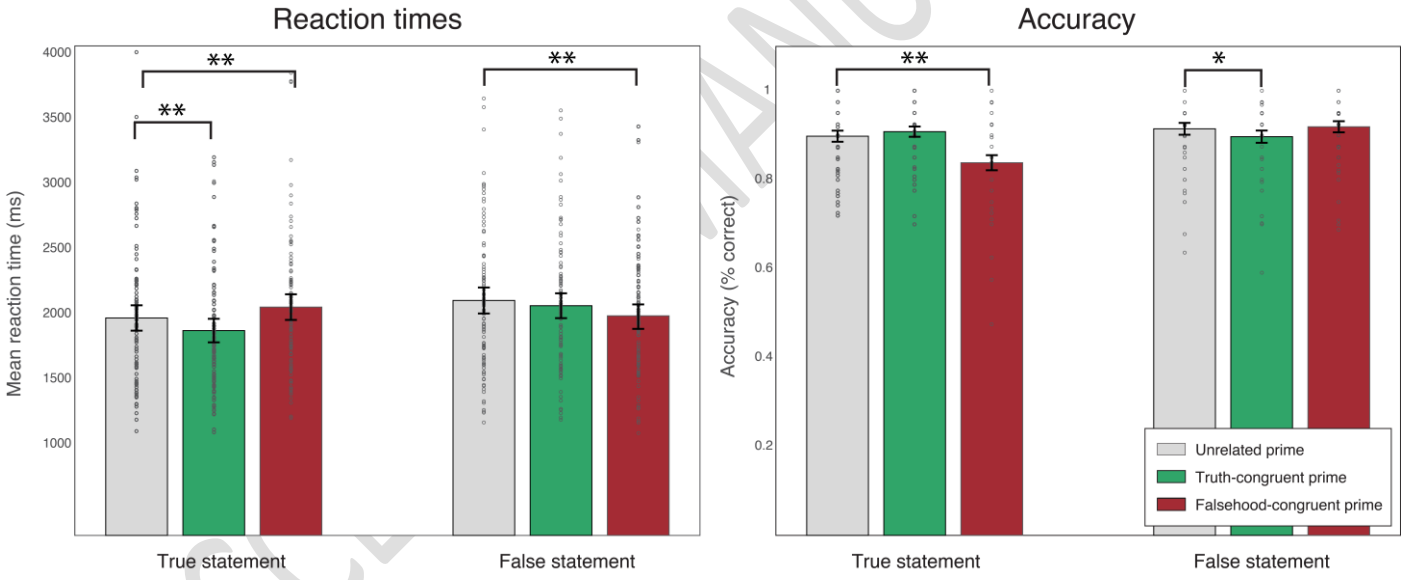

Figure 4: Mean reaction times (left) and accuracy (right) of true/false judgment as a function of Sentence Validity and Prime Congruency in Experiment 3. Error bars depict 95% confidence intervals. Circles denote data points of individual subjects. Significant contrasts are indicated (** equals $p \leq .001$; * equals $p < .05$).

4.2.2. Accuracy. We fitted a model including Sentence Validity, Prime Congruency and the two-way interaction of Sentence Validity and Prime Congruency as fixed effects with by-subject and by-item random intercepts. The analysis revealed a main effect of Sentence

Validity, $\chi^2(1) = 65.38$, $p < .001$, a main effect of Prime Congruency, $\chi^2(2) = 45.04$, $p < .001$, and a two-way interaction of Sentence Validity and Prime Congruency, $\chi^2(2) = 100.81$, $p < .001$ (see Figure 4, right). The simple contrast showed that, for true sentences, participants were less accurate following falsehood-congruent primes ($M_{prob} = .87$, $SE = .01$), as compared to unrelated primes ($M_{prob} = .93$, $SE = .01$). There was no difference in accuracy between unrelated and truth-congruent primes ($M_{prob} = .94$, $SE = .01$; see Table 8). For false sentences, participants were less accurate following truth-congruent primes ($M_{prob} = .93$, $SE = .01$), as compared to unrelated primes ($M_{prob} = .94$, $SE = .01$). There was no difference in accuracy between unrelated and falsehood-congruent primes ($M_{prob} = .95$, $SE = .01$; see Table 8; see also Table 9).

Table 8: Comparisons of log odds ratio of participants' accuracy, p-values are adjusted by Tukey-method, Experiment 3.

Contrast	*Odds Ratio*	*Standard Error*	*p-value*
True Sentences			
Unrelated vs. truth-congruent prime	.87	.07	.17
Unrelated vs. falsehood-congruent prime	1.84	.13	< .001
False Sentences			
Unrelated vs. truth-congruent prime	1.29	.10	.004
Unrelated vs. falsehood-congruent prime	.94	.08	.71

Table 9: Bayes Factors for main effects, interactions and simple comparisons of participants' reaction times and accuracy, Experiment 3.

	Bayes Factors	
Effects and Simple Comparisons	*Reaction Times*	*Accuracy*
Sentence Validity	BF_{10} = 775758.08	BF_{10} = 957169.33
Prime Congruency	BF_{10} = 19.97	BF_{10} = 1451.24
Sentences Validity × Prime Congruency	BF_{10} = 8.828e+10	BF_{10} = 4.410e+12
True sentences		
Unrelated vs. truth-congruent prime	BF_{10} = 363.95	BF_{10} = .36
Unrelated vs. falsehood-congruent prime	BF_{10} = 123.63	BF_{10} = 5.876e+6
False sentences		
Unrelated vas. truth-congruent prime	BF_{10} = 1.32	BF_{10} = 8.66
Unrelated vs. falsehood-congruent prime	BF_{10} = 181632.93	BF_{10} = .21

4.3. Discussion

Experiment 3 demonstrated again the facilitating effect of pre-activating background knowledge on validation performance. As for the differential effects of knowledge types, validation was again facilitated by primes that were congruent with the content of the sentence, irrespective of sentences' validity. Replicating the previous experiments, Experiment 3 showed that truth-congruent primes only led to facilitation of judging true sentences, while falsehood-congruent primes slowed down reaction times for true sentences but led to facilitation of judging false sentences. The error pattern of Experiment 3 showed that falsehood-congruent primes reduced accuracy for true sentences and truth-congruent primes reduced accuracy of false sentences. Thus, in Experiment 3, both the reaction times as well as the error pattern suggest that performance was mainly driven by conceptual overlap between the primes and the sentences. And, importantly, no facilitative effect for false

sentences was found for truth-congruent primes. Taken together, these results imply that comprehenders rely on a mismatch detection between the sentence components via activation of semantic associations, and do not benefit from the activation of truth-congruent information (e.g., Ferreira, Bailey, & Ferraro, 2002).

But is this facilitation unique to information that determines the truth value of the sentence, or is it obtained by the activation of any sentence component? In the present study, although truth-congruent primes naturally share semantic overlap with both the subjects and objects of sentences (e.g., 'tracks' share semantic overlap with 'trains'), they did not seem to facilitate comprehension of false sentences (that is, an image of 'tracks' did not facilitate validity judgments of the sentence 'highways are the infrastructure trains run on/trains run on highways', compared to the unrelated prime (i.e., an image of a TV test pattern). On the contrary, they seemed to cause more judgment errors. Notably, errors were present in Experiment 2 and 3 but not in Experiment 1, suggesting that the proximity between prime and false concept might be a factor of influence. The error pattern might be explained by a conflicting activation of the primed concepts (e.g., 'tracks') and the presented concept at the beginning of the sentence (e.g. 'highways'), potentially hampering comprehension. Thus, Experiment 4 investigated the relative activation of sentence components by primes that are either identical (i.e., sentence-congruent) or associatively related and also truth-related (i.e., truth-congruent) with respect to a particular sentence, eliminating influences of conflicting activations.

5. Experiment 4

Experiment 4 had two main goals. The first goal was to investigate the role of relative activation of sentence components by sentence-congruent and truth-congruent primes for validation. To allow for a meaningful comparison between sentence-congruent and truth-congruent primes, we kept the distance between prime and to-be-activated component

constant across the two priming conditions for false sentences. Moreover, none of the primes created a conflict with the beginning of sentences. The second goal was to explore whether the effects found in Experiment 1-3 are unique to picture primes, or evoked by word primes as well, suggesting they tap onto a more general mechanism of comprehension. An additional, related goal, was to better control for the associations participants might have to the presented primes. For example, in Experiment 1-3 we presented a picture of a circle as prime, to activate the concept 'round' prior to the sentence 'A tennis ball is square/Square is the shape of a tennis ball'. But a circle could also activate other concepts (e.g., circle, ball) rather than the exact one we were aiming for. By using words, we were able to directly activate the relevant concepts.

To this end, participants were presented with true and false sentences (e.g., 'trains run on tracks/highways') and primed either with the subject of the sentence (e.g., 'trains'; sentence-congruent prime) or with the true concept related to a sentence (e.g., 'tracks'; truth-congruent prime). For true sentences, we predicted the fastest true judgments for sentence-congruent primes (identical to the subject of the sentence, associatively related to the true concept presented at the end of the sentence), followed by truth-congruent primes (identical to the true concept presented at the end of the sentence, associatively related to the subject of sentence). Thus, we expected the strongest priming effects for primes identical with sentence components and presented in close proximity to the to-be-activated-component. Both types of primes should lead to faster reactions as compared to unrelated primes. For false sentences, we predicted the fastest false judgments for sentence-congruent primes (identical to the subject of the sentence), followed by truth-congruent primes (associatively related to the subject of the sentence). Both types of primes should lead to faster reactions as compared to unrelated primes.

5.1. Methods

5.1.1. Participants and design. We recruited 100 participants (46 female, 54 male; M_{age} = 34.59; SD_{age} = 12.71) via Prolific Academic (see Palan & Schitter, 2018; Peer et al., 2017). Participants could only sign up for the experiment if they resided in the United States and were native English speakers. We required that they had previously completed at least 10 tests via Prolific Academic and held a record of supplying acceptable data at least 95% of the time. They received £1.67 (approx. $2) for their participation. The post-experimental questionnaire was identical to previous experiments. The study consisted of a 2 (Sentence Validity: true vs. false) × 3 (Prime Congruency: truth-congruent vs. sentence-congruent vs. unrelated) within-participants design.

5.1.2. Procedure and Stimuli. The procedure was largely identical to previous experiments with the following exceptions: first, none of the sentences contained fictional elements, to avoid a potential influence on validation by unrealistic content. Second, we ensured that false versions of sentences did not include concepts that are direct opposites of the true concepts for the sentence (e.g., 'Peas are triangular' instead of 'Peas are square'; see Appendix C). The mean number of words per sentence was 4.07 (*SD* = 1.07). All sentences had the same general structure, mentioning the subject before the concept (similar to Experiment 1). The mean number of words between prime and concept was .27 (*SD* = .44).

Third, and most importantly, instead of pictures, words were used as primes to reduce any ambiguity that might have been caused by metaphorical representations in previous experiments. Words were either identical with the subject of the sentence (e.g., 'peas' before 'peas are round/triangular) or truth-congruent (e.g., 'round' before 'peas are round/triangular). Unrelated primes were meaningless letter strings (i.e. 'xxxx'). Each trial started with a warning signal (+++), presented in the center of the screen for 500 ms. Subsequently, a word was presented for 200 ms, followed by a true or false sentence, separated by a 50 ms blank screen (see Figure 5). The sentence stayed on the screen until participants indicated their answer. Each trial was separated by a 1000 ms interval.

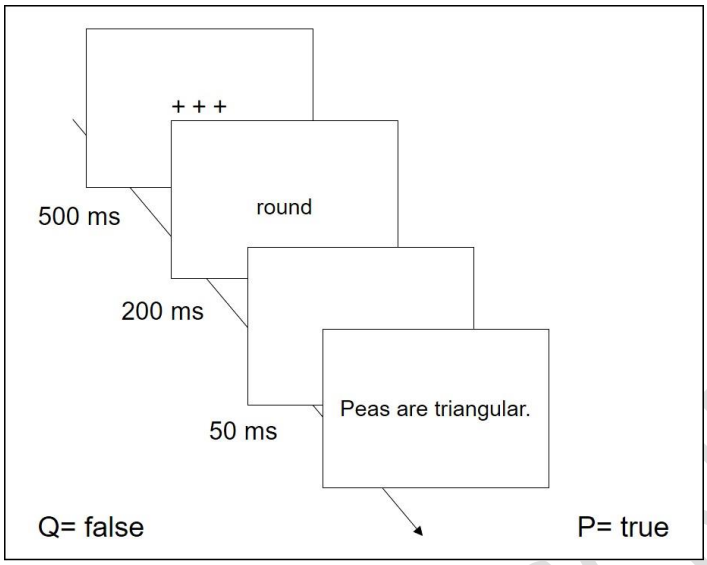

Figure 5: Example of trial sequence in Experiment 4

5.2. Results

2% of participants reported being non-native English speakers and 6% were in the presence of others while performing the task. These participants are included in the following analyses. To investigate whether results are affected, we excluded these participants in a separate analysis. Results show that the general pattern of effects stays the same. We only included trials with reaction times above 300 ms and below 10000 ms (99.2% of trials). The following analyses are based on 100 participants.

5.2.1. Reaction times. We excluded all trials in which sentences were judged incorrectly (6.9% of trials). To test whether pre-activating truth- and sentence-congruent concepts influenced reaction times of validity-judgments, we fitted a model including Sentence Validity, Prime Congruency and the two-way interaction of Sentence Validity and Prime Congruency as fixed effects. The model featured by-subject and by-item random intercepts, as well as a by-subject and by-item random slopes for Sentence Validity. The analysis showed a main effect for Sentence Validity, $\chi^2(1) = 53.78$, $p < .001$, a main effect for Prime

Congruency, $\chi^2(2) = 217.67$, $p < .001$, and a two-way interaction of Sentence Validity and Prime Congruency, $\chi^2(2) = 29.78$, $p < .001$ (see Figure 6, left). Simple contrast showed that for true sentences participants judged sentences faster when they were preceded by sentence-congruent primes (M_{EM} = 1299.03, SE = 48.25) and by truth-congruent primes (M_{EM} = 1391.29, SE = 48.26) as compared to unrelated primes (M_{EM} = 1470.30, SE = 48.26). As predicted, sentence-congruent primes led to more facilitation compared to truth-congruent primes (see Table 10). For false sentences, participants judged sentences faster when they were preceded by sentence-congruent primes (M_{EM} = 1407.60, SE = 51.20) as compared to unrelated primes (M_{EM} = 1574.39, SE = 51.20). Contrary to our prediction, truth-congruent primes (M_{EM} = 1615.95, SE = 51.17) did not lead to facilitation. On the contrary, they showed the tendency to slow down reactions compared to unrelated primes (see Table 10). Thus, sentence-congruent primes facilitated reactions for both true and false sentences, while truth-congruent-primes only facilitated the judgment of true sentences (see also Table 12).

Table 10: Comparisons of estimated marginal means of participants' reaction times, p-values are adjusted by Tukey-method, Experiment 4.

Contrast	*Estimate*	*Standard Error*	*p-value*
True Sentences			
Unrelated vs. sentence-congruent prime	171.27	17.89	< .001
Unrelated vs. truth-congruent prime	79.01	17.73	< .001
Sentence-congruent vs. truth-congruent	-92.26	17.69	< .001
False Sentences			
Unrelated vs. sentence-congruent prime	166.79	17.78	< .001
Unrelated vs. truth-congruent prime	-41.56	17.71	.05

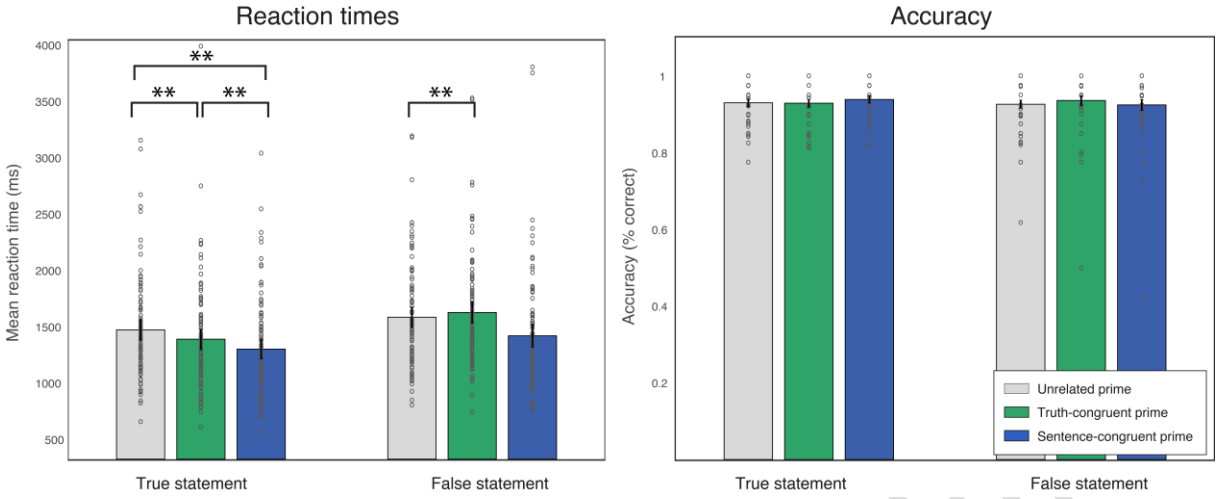

Figure 6: Mean reaction times (left) and accuracy (right) of true/false judgment as a function of Sentence Validity and Prime Congruency in Experiment 4. Error bars depict 95% confidence intervals. Circles denote data points of individual subjects. Significant contrasts are indicated (** equals $p \leq .001$; * equals $p < .05$).

5.2.2. Accuracy. We fitted a model including Sentence Validity, Prime Congruency and the two-way interaction of Sentence Validity and Prime Congruency as fixed effects with by-subject and by-item random intercepts. The analysis revealed a marginally significant main effect of Sentence Validity, $\chi^2(1) = 3.56$, $p = .059$, and a two-way interaction of Sentence Validity and Prime Congruency, $\chi^2(2) = 8.28$, $p = .016$ (see Figure 6, right). The simple contrast showed that, for true sentences, participants' accuracy did not differ following sentence-congruent primes ($M_{prob} = .97$, $SE = .004$), truth-congruent primes ($M_{prob} = .96$, $SE = .004$) or unrelated primes ($M_{prob} = .96$, $SE = .004$; see Table 11). Also for false sentences, accuracy did not differ following sentence-congruent primes ($M_{prob} = .96$, $SE = .005$) and truth-congruent primes ($M_{prob} = .96$, $SE = .004$) compared to unrelated primes ($M_{prob} = .96$, $SE = .005$; see Table 11; see also Table 12).

Table 11: Comparisons of log odds ratio of participants' accuracy, p-values are adjusted by Tukey-method, Experiment 4.

Contrast	*Odds Ratio*	*Standard Error*	*p-value*
True Sentences			
Unrelated vs. sentence-congruent prime	.84	.08	.16
Unrelated vs. truth-congruent prime	1.01	.09	.99
False Sentences			
Unrelated vs. sentence-congruent prime	1.02	.09	.98
Unrelated vs. truth-congruent prime	.83	.08	.11

Table 12: Bayes Factors for main effects, interactions and simple comparisons of participants' reaction times and accuracy, Experiment 4.

	Bayes Factors	
Effects and Simple Comparisons	*Reaction Times*	*Accuracy*
Sentence Validity	BF_{10} = 1.585e+24	BF_{10} = .13
Prime Congruency	BF_{10} = 6.547e+20	BF_{10} = .03
Sentences Validity × Prime Congruency	BF_{10} = 3190.89	BF_{10} = .28
True sentences		
Unrelated vs. sentence-congruent prime	BF_{10} = 6.576e+12	BF_{10} = .61
Unrelated vs. truth-congruent prime	BF_{10} = 411.87	BF_{10} = .15
Sentence-congruent vs. truth-congruent	BF_{10} = 216.82	
False sentences		
Unrelated vas. sentence-congruent prime	BF_{10} = 4.047e+9	BF_{10} = .16
Unrelated vs. truth-congruent prime	BF_{10} = 2.94	BF_{10} = .73

5.3. Discussion

Experiment 4 demonstrated that activating content that is part of true and false sentences facilitates validity-judgments. This finding is in line with the assumption that validation relies on knowledge about semantic network affiliations. Interestingly, and differently from our prediction, truth-congruent primes tended to interfere with validity-judgments for false sentences, despite the semantic overlap with the subject of a sentence. Even if this marginally significant (and close to conclusive) result is not true, it is clear that truth-congruent primes did not facilitate validation. This is surprising, as it does not only go against the claim that activation of background knowledge aids validation; it is also incompatible with the widely-accepted claim that priming a semantically related concept (e.g., 'tracks' for 'trains') facilitates sentence comprehension (cf. Meyer & Schvaneveldt, 1971; Neely, 1977; Posner & Snyder, 1975). Arguably, the conflict between the truth-congruent concept and the last word of the sentence that makes it false (e.g., 'highways') could have overshadowed the initial facilitation of the first word, that is expected given previous literature. Thus, it seems like activating truth-congruent concepts might even tamper with validation processes, given the evoked conflict with the actual content of the sentence.

Notably, the latter finding further demonstrates that although the relations between the prime and the subject of the sentence strongly affect participants' performance (hence, the greater facilitation of processing true sentences following a prime that is identical to the first rather than last word), they do not fully explain participants' behavior. Indeed, although truth-congruent primes were naturally well-associated with the subject of the sentence, their conflict with the last word abolished their facilitative effect. Thus, a pre-activation of the subject of the sentence was not the sole factor affecting validity-judgments.

6. General Discussion

The aim of the present study was to assess the effect of knowledge pre-activation on validation processes and to investigate whether different types of knowledge exert a different effect on these processes. While it seems beyond debate that existing knowledge plays an important role for validation processes (Singer, 2019), the effects of pre-activating such knowledge on explicit validation processes have not been tested. In addition, it seems less clear whether people need to compare presented information with the correct answer (i.e., true concept) to determine that a sentence is false, or whether they rely on a mismatch between presented sentence components (i.e., detect semantic network affiliation) in order to spot falsehood. The results of our study support the latter interpretation, showing no facilitation of validation processes (and perhaps even a negative effect) for activated true concepts.

We pre-activated participants' knowledge by priming concepts that are either congruent with true concepts, congruent with false concepts or congruent with specific components of the sentences and measured speed and accuracy of validity-judgments. The results of Experiment 1 suggested that truth-congruent primes speed up validation of true sentences and falsehood-congruent primes speed up validation of false sentences. Thus, the results of Experiment 1 suggest that validation is carried out on the basis of knowledge about semantic network affiliation. Experiment 2 used a different sentence structure, with concepts presented at the beginning rather than the end of sentences, and replicated the result pattern of Experiment 1, corroborating the robustness of the effect. Yet, the error pattern in Experiment 2 was neither in line with semantic integration accounts (e.g., Berkum, Hagoort, & Brown, 1999), nor with models of validation (e.g., O'Brien, & Cook, 2016; Richter et al. 2009). Surprisingly, for false sentences, participants made more errors following both truth-congruent and falsehood-congruent primes. Importantly, however, this surprising effect was not replicated in Experiment 3, suggesting that it might have been a false-positive result. In Experiment 3, truth-congruent primes again led to facilitation of true judgments and falsehood-congruent primes led to the facilitation of false judgements, replicating the results

from Experiment 1 and 2. Moreover, the error pattern showed that falsehood-congruent primes reduced accuracy for true sentences and truth-congruent primes reduced accuracy of false sentences. The facilitation effects and the error pattern of congruent primes (falsehood-congruent for false sentences and truth-congruent for true sentences) clearly imply that validation can be carried out with minimal semantic processing that is just good enough to detect semantic network affiliation (Ferreira, Bailey, & Ferraro, 2002).

Finally, Experiment 4 showed that activating a specific component of true and false sentences facilitates validity-judgments, but to a different degree. Importantly, and contrary to our expectations, pre-activating truth-congruent concepts did not facilitate falsehood-judgments, despite being semantically or associatively related to the subject of the sentence. To illustrate, we expected that pre-activating 'tracks' will activate 'train' to some degree. However, the results of Experiment 4 suggest that such activation is not sufficient to speed up the validation of false sentences, and might even tamper with the process, showing a marginally significant decrease of performance for truth-congruent primes. Speculatively, this lack of facilitation (and possible semantic interference) following truth-congruent primes stems from the conflict between these prime and the actual content of the false sentence (e.g., Piai, Roelofs, & van der Meij, 2012).

Taken together, our results strongly support the claim that validation can be carried out with minimal semantic processing that is just good enough to detect a lack of semantic overlap, or a mismatch, between sentence components, and does not benefit from pre-activating the specific true concept for a false sentence (see Figure 7). Activating content-congruent concepts facilitated validation for true and false sentences. Moreover, activating incongruent concepts led to more errors in judging the validity of true and false sentences. Thus, background knowledge needs not involve the activation of the true content of the sentence, but rather simply the association of the sentence's content to a specific sematic network. Put differently, it is sufficient to know that 'macadamia are not berries' or that 'the

category of berries does not include macadamia' for validation to take place. Knowing that 'macademia are nuts' does not seem to help validation, and might even interfere with it. Under this account, when encountered by the sentence 'macadamia are berries', the validation process only entails activating the two concepts (i.e., 'macadamia' and 'berries'), and detecting the lack of overlap in features (in terms of distributed models; see Masson, 1995; Moss, Hare, Day, & Tyler, 1994), or the lack of connection between the units (in line with spreading activation models; Anderson, 1983; Collins & Loftus, 1975). And so, pre-activating 'berries' or 'macadamia' facilitates the validation process, as one of the two critical components of the sentence has already been activated, and the search for a match and detecting a mismatch should accordingly be easier.

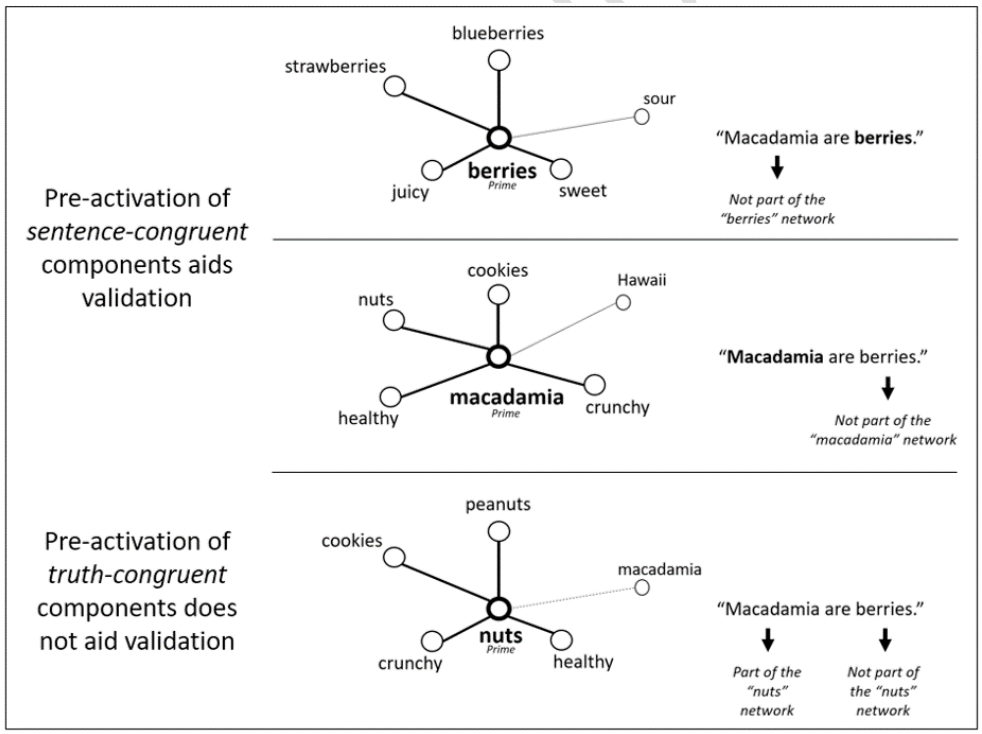

Figure 7: Schematic overview of the validation mechanism suggested by the results.

Within this framework, our findings imply that background knowledge has an essential role in the validation process, in line with contemporary models of validation (Cook & O'Brien, 2014; O'Brien & Cook, 2016; Richter, 2015; Singer, 2013). The RI-Val model

(O'Brien, & Cook, 2016) postulates a validation stage, in which active linked components in short-term memory are validated against contents from long-term memory by a pattern-matching process. Our findings suggest that the relevant knowledge that is required for this stage is that linked sentence components cannot be integrated, rather than retrieving the correct answer from long-term memory. That is, rather than detecting the mismatch between the active link 'macadamia – berries' and the link from long-term memory 'macadamia – nuts,' the mismatch between 'macadamia' and 'berries' is detected. Accordingly, determining the relations between a sentences' components may be done in a low-level manner, relying on statistical co-occurrences and category matching (Pulvermüller, 2013) not on expected content.

This interpretation is also in line with the Good-Enough Representations approach in language processing (Ferreira, Bailey, & Ferraro, 2002). This approach suggests that semantic representations might be just good enough to process and comprehend information, depending on the task that a comprehender has to perform. Arguably, comprehenders do not always engage in complete and detailed processing of the sentence, which only occurs when needed. And so, activating content-congruent concepts might just be enough for validation. Yet, if validity of a sentence cannot be determined due to semantic network affiliation (e.g., Erickson, & Mattson, 1981; Sanford, 2002) validation might require a different strategy (e.g., Cook, Walsh, Bills, Kircher, & O'Brien, 2016) and speculatively, the activation of true concepts might be helpful under such circumstances.

How do our findings relate to accounts assuming that validation is influenced by the perceived fluency of information processing (e.g., Reber & Schwarz, 1999)? Experienced ease of processing (i.e. fluency) is typically associated with truth (see Unkelbach, 2007), so that when the processing of information is experienced as fluent, it might be judged as true even if it is not (see Lewandowsky, Ecker, Seifert, Schwarz, & Cook, 2012). Accordingly, when congruent content of a sentence is activated, it should increase perceived fluency and

lead to the judgment that a sentence is true. On the other hand, if incongruent content is activated, it should interfere with sentence processing, given the resulting conflict. This then decreases perceived fluency and may lead to the judgment that the sentence is false (see Newman et al., 2015; cf. Hansen, Dechene, & Wänke, 2008). Our results, however, suggest that greater fluency, which presumably follows the activation of a content-congruent prime, might possibly lead not only to faster true but also faster false judgments, as opposed to the above approach. A possible explanation to this apparent contradiction might lie in the fact that our sentences were easy to validate (i.e., they were clearly false/true), while fluency is assumed to influence judgments under uncertainty (e.g., Jacoby & Kelley, 1987; Zajonc, 1968; see also Loersch & Payne, 2011), that is, when background knowledge is not available or less accessible, which was not the case in our experiments.

A similar argument can be made with respect to the finding that the error pattern we obtained stemmed from incongruent primes reducing accuracy of judgments (compared to unrelated primes), rather than from congruent primes improving accuracy of judgments. This again may have to do with the fact that we used sentences that were easy to validate, suggesting a potential ceiling-effect, leaving not much room for improvement. In a way, this echoes the finding that incongruent scene-context hinders performance but congruent one does not improve it, compared with a 'no context' condition (Davenport & Potter, 2004). Yet, when using ambiguous, hard to detect objects, contextual facilitation by congruent objects is found (Brandman & Peelen, 2017). Future research is needed to better clarify this point.

Three alternative explanations of our results should be considered. First, a potential worry in priming experiments is that primes did not activate concepts to the same degree across conditions (e.g., in our case, the relation between 'stopwatch' and 'fast' might be stronger/weaker than the relation between 'hourglass' and 'slow'). Yet, this concern is not applicable to our design, because a single nominal prime appeared before four nominal sentences (e.g., 'stopwatch' and 'hourglass' appeared before 'Cheetahs run fast' and

'Cheetahs run slow' and before 'Turtles move slow' and 'Turtles move fast'; counterbalanced across participants,). Thus, each nominal prime could either serve as a congruent or an incongruent prime, across participants. Still, as some primes did not refer to literal but metaphorical representations, such metaphorical primes might have induced weaker priming effects across conditions. However, excluding these metaphorical primes did not change the pattern of results. Importantly, all of these concerns are addressed by the use of words in Experiment 4, hereby making the primes less open to interpretation.

A second alternative explanation concerns the fact that for Experiments 1-3, out of 240 sentences used in our experiments, six contained information related to unreal or fictional events (see Appendix A and B). Thus, the inclusion of these sentences might have influenced general processing tendencies that might have affected the results in our study. For example, in a fictional world, turtles might be fast, thus activating the concept 'fast' might be considered activating a truth-congruent rather than a falsehood-congruent concept. Yet, this would imply that participants tended to judge false sentences as true. Performance for fictional sentences was not different than for the other, non-fictional ones (90.7% overall accuracy for fictional sentences and 90.8% for non-fictional ones). Moreover, the fictional sentences referred to well-known fictional facts (e.g., Santa Claus brings presents on Christmas), thus answering them correctly required access to general fictional knowledge, whereas imaging the possibility that there might be fast turtles in a fictional world, might involve different cognitive mechanisms. And, perhaps most convincingly, these sentences were not included in Experiment 4 (see Appendix C). Nevertheless, future research should investigate whether an unrealistic context in particular (Rapp, Hinze, Slaten, & Horton, 2014) or putting participants in a certain mindset in general (e.g., Mayo, 2015; Schul, Mayo, & Burnstein, 2004) might moderate the present findings.

A third alternative explanation for our findings should be considered with respect to models that assume that comprehension and validation are separate stages of information

processing (e.g., Connell & Keane, 2006; Gilbert, 1991). Connell and Keane (2006) propose a two-step process in which incoming information is first comprehended by mentally representing the information together with related prior knowledge in a comprehension stage. Then, at the second stage the information is evaluated by comparing whether activated mental representations fit with prior knowledge in an assessment stage. To the extent that comprehension and validation form two separate stages of information processing, our results might simply reflect lexical activation and thus, facilitate comprehension, rather than validation. And so, because priming sentence congruent content facilitates comprehension, the subsequent stage of validation can also be carried out faster. We, on the other hand, derived our hypotheses from the theory that validation is a by-product of comprehension, which seems more parsimonious (Cook & O'Brien, 2014; O'Brien & Cook, 2016; Richter, 2015; Singer, 2013). That is, co-activated prior knowledge, necessary to comprehend incoming information, is simultaneously used to validate this information (Richter, 2015). Yet, we acknowledge that our experiments and results alone do not allow to determine whether the pre-activation of sentence-congruent content aids validation directly (i.e., validation as by-product of comprehension) or indirectly (i.e., comprehension followed by validation). Future studies should address this question by comparing the facilitating effect of sentence-congruent primes between a validation task (i.e., determining the truth-value of sentences) and a comprehension task (i.e., asking questions about sentence content; see Isberner & Richter, 2014). No differences in reaction times between the two tasks should be predicted if validation happens as a by-product of comprehension, but reaction times for the comprehension task should be faster if validation follows comprehension.

A final concern relates to the generalizability of our findings to everyday validation procedures. Our paradigm required participants to explicitly judge the validity of sentences (while in reality individuals are typically not explicitly asked to judge whether information is true of false). It could be argued that explicit validity-judgments make strategic processes

more likely and so participants might try to optimize strategies to complete the task as fast and accurate as possible. Accordingly, participants were more likely to use knowledge about semantic network affiliation in our study while they might be more likely to use true concepts under different conditions (see Richter & Maier, 2017). Our present findings cannot rule out this possibility and future research should investigate under which conditions people rely on a mismatch detection between sentence components and under which conditions they use true concepts to validate information. Still, our findings clearly show that validation *can* be carried out only on the basis of detection of lack of semantic overlap between sentence components and does not need to involve the activation of true concepts. Evidently, a precondition for validation based on knowledge about semantic network affiliation is that this criterion is sufficient to detect a mismatch between sentence components (as was the case for the sentences used in our paradigm). Thus, our findings are generalizable to information for which this is the case. Yet, speculatively, information for which this is not the case (e.g., 'cows drink milk') might require different validation processes.

How do our findings inform previous research that demonstrated the error-proneness of validation processes and their effect on memory (for an overview, see Rapp & Braasch, 2014)? The surprising finding that people rely on inaccurate information despite knowing that the information is false (e.g., Eslick, Fazio, & Marsh, 2011; Fazio, Barber, Rajaram, Ornstein, & Marsh, 2013; Marsh & Fazio, 2006; Marsh, Meade, & Roediger, 2003), might be explained by decreased activation of truth-congruent concepts and increased activation of falsehood-congruent concepts. If people indeed do not access the correct answer while detecting falsehood, correct knowledge cannot compete with the memory trace that might be left by false information. Accordingly, this trace might increase reliance on false information (see Rapp, 2016; Rapp, Hinze, Kohlhepp, & Ryskin, 2014). Our present findings imply that validation can be carried out on the basis of knowledge about semantic network affiliation. Accordingly, during information processing, present sentence components might receive the

strongest activation, while true knowledge is weaker or not activated at all. Consequently, strongly activated components have a high likelihood to be encoded in memory and lead to false information being remembered despite being detected as false (Weil, Schul & Mayo, 2019).

6. Conclusion

The present study demonstrates that during validation of false sentences, people do not benefit from access to the correct answer. It might be enough to know that something is *not* the case in order to detect falsehood. And so, when confronted with a false sentence, like 'macadamia are berries', one would not benefit from being reminded that macadamia are nuts. On the contrary, our findings imply that bringing the truth to mind (i.e., activating truth-congruent concepts) before being confronted with falsehood might even be detrimental, leading to errors in judgments and possibly slower judgment times. The findings might also shed fresh light on previous findings that demonstrated the error-proneness of validation processes and their effect on memory: the detection of falsehood does not require bringing the truth to mind. Thus, even if false information is identified as false, in the absence of a competing truth-congruent concept people might rely on inaccurate information.

Acknowledgements

This work was supported by a research fund from the University of Hull, UK, and by the Israel Science Foundation [grant number 1847/16].

Supplementary material

Materials and data are available at https://osf.io/c6j4b/.

APPENDIX A

Overview of true versions of sentences used in Experiment 1. False versions were created by switching the last word (highlighted) of Set A and Set B for each corresponding sentence.

Set A	**Set B**
The sun rises in the **east**.	The sun sets in the **west**.
Turtles move **slow**.	Cheetahs run **fast**.
Giraffes are **tall**.	Dwarfs are **short**.
A forest consists of **trees**.	On the beach, kids build castles out of **sand**.
Scissors are used to **cut**.	Glue is used to **paste**.
England is a **country**.	Berlin is a **city**.
Boats sail on **water**.	Horses run on **land**.
Shorts are worn in **summer**.	Gloves are worn in **winter**.
Sugar is **sweet**.	Beer is **bitter**.
Lemons are **sour**.	Potato chips are **salty**.
The Arctic is **cold**.	A fire is **hot**.
Fish have **scales**.	Humans have **pores**.
Birds have **feathers**.	Dogs have **fur**.
Wine is **liquid**.	Walls are **solid**.
Soap makes you **clean**.	Grease makes you **dirty**.
Towels are used to get **dry**.	Rain makes you **wet**.
Cars have **wheels**.	Gazelles have **legs**.
Easter eggs are hidden by **bunnies**.	Roosters are male **chicken**.
The Statue of Liberty is in **New York**.	The Eiffel tower is in **Paris**.
John Lennon demonstrated for **peace**.	In 1939 the world was at **war**.
You should cross the road when your traffic light is **green**.	Blood is **red**.
Flashlights are **bright**.	When you switch off the light it is **dark**.
Elephants are **big**.	Ants are **small**.
Airplanes fly in the **sky**.	Cars drive on the **street**.
Snow is **white**.	A panther is **black**.
Rocks are **hard**.	Silk is **soft**.
Jumbo jets are **heavy**.	Feathers are **light**.
A pear is a **fruit**.	Broccoli is a **vegetable**.
Baked beans is a **dish**.	Lemonade is a **drink**.
Diamonds are **expensive**.	Discounters are **cheap**.
Owls are active during the **night**.	Kids go to school during the **day**.
Peas are **round**.	A TV screen is **square**.
A year has 12 **months**.	A day has 24 **hours**.
Magazines are made from **paper**.	Clothes are made from **textile**.
You lock a door with a **key**.	You cut bread with a **knife**.
During a theatre performance the audience should be **silent**.	Rock concerts are **loud**.
The lowest story of a building is the **basement**.	If you take the stairs up you reach the **attic**.
Cake is baked in the **oven**.	Ice cream is kept in the **fridge**.
Fever, coughing and a running nose are signs that you are **sick**.	You leave the hospital when you are **healthy**.
Winning the lottery is **rare**.	Preferring holidays over workdays is **common**.
Trains run on **tracks**.	Buses run on **highways**.
Honey is made by **bees**.	The insects, developing from caterpillars are **butterflies**.

Most people´s dominant hand is **right**.
You need to charge your battery when it is **empty**.
Texas lies in the **south**.
A chair is a piece of **furniture**.
Alligators are **reptiles**.
An eye of a needle is **narrow**.
The currency in the USA is **dollar**.
Dogs wag their **tails**.
Health, luck and enough money makes people **happy**.
Breakfast is eaten in the **morning**.
Soup is eaten with a **spoon**.
You enter your house through the **door**.
People lie down to sleep in the **bedroom**.
You chew with your **teeth**.
People drink coffee from a **cup**.
Scarves are worn around the **neck**.
Shoes are worn on **feet**.
Tom Hanks is an **actor**.
Tuna is sold in **cans**.
Four is a **number**.
Tigers have **paws**.
Moles live **underground**.
Violins have **strings**.
Scrapers are **DIY tools**.
Ostriches are **bipeds**.
Macadamia are **nuts**.
A tennis ball is **round**.
A cow gives **milk**.
Kids play 'trick or treat' on **Halloween**.
Roses have **thorns**.
Bees live in **hives**.
Rye belongs to the family of **grains**.
You smell with your **nose**.
Japan is in **Asia**.
In a garage people fix **vehicles**.
The USA is a **democracy**.
Van Gogh was a **painter**.
Johnnie Walker is a brand of **whiskey**.
Cabernet Sauvignon is a name for **wine**.
Pregnancy happens in the bodies of **women**.
The sky is **blue**.
Neil Armstrong flew to the **moon**.
Barcelona is in **Spain**.
Rolex produces **watches**.
You close a jacket with a **zipper**.
Louis Armstrong played the **trumpet**.
A witch rides her flying **broom**.
Cinderella’s carriage was made from **pumpkin**.

In Britain people drive on the **left**.
The bath tub is overflowing when it is **full**.
Alaska lies in the **north**.
A sweater is a piece of **clothing**.
Wolves are **mammals**.
A baseball field is **wide**.
The currency in Russia is **ruble**.
Left and right from its nose, a cat has **whiskers**.
The death of a dear one makes people **sad**.
Dinner is eaten in the **evening**.
Salad is eaten with a **fork**.
To let air inside the car while driving, you open the **window**.
Dinner is cooked in the **kitchen**.
The doctor says: 'Stick out your **tongue**!'
You fry a steak in a **pan**.
Watches are worn around the **wrist**.
Wedding rings are worn on the left **hand**.
Bob Marley was a **singer**.
Water is sold in **bottles**.
Y is a **letter**.
Children have **hands**.
Squirrels live **on trees**.
Accordions have **keyboards**.
Harps are musical **instruments**.
Horses are **quadrupeds**.
Black currants are **berries**.
Dice are **square**.
If you squeeze oranges you get **juice**.
Santa Claus brings presents on **Christmas**.
Hedgehogs have **stings**.
Bears live in **caves**.
Seaweed belongs to the family of **algae**.
You hear with your **ears**.
The Netherlands are in **Europe**.
With clay and a potter's wheel you create **pottery**.
Saudi Arabia is a **monarchy**.
Shakespeare was a **poet**.
Pepsi is a brand of **soda**.
Absolut and Smirnoff are brands of **vodka**.
Beard growth is a sign of puberty in **men**.
Corn is **yellow**.
The red planet is **Mars**.
Hamburg is in **Germany**.
Macs from Apple are **computers**.
You open a door with a **handle**.
Jimi Hendrix played the **guitar**.
Aladdin rides the magic **carpet**.
For strength Popeye eats **spinach**.

Orangutans are **primates**.
Kobe Bryant is a **basketball player**.
A sandbox is a playground for **children**.
The alphabet consists of **letters**.
A birth date consists of **numbers**.
Cobras are **snakes**.
A black widow is a **spider**.
Lipsticks are **cosmetics**.
Chips are made from **potatoes**.
Mirrors are made from **glass**.
Beech is a **wood**.
Salami is a **sausage**.
A shower can be found in the **bathroom**.
Brooms are used to sweep the **floor**.
A pregnancy takes 9 **months**.
Candles work with **fire**.
Sunglasses protect your **eyes**.
Coffee is made from **beans**.
Sheep live on **land**.
Hamsters have **fur**.
Mud makes you **dirty**.
Swimming makes you **wet**.
Balloons are **light**.
A cucumber is a **vegetable**.
Pants are made from **textile**.
You cut wood with a **saw**.
You warm food in the **microwave**.
A day ends in the **evening**.
You leave the house **in shoes**.
You wash your hands with **soap**.

Beavers are **rodents**.
Cristiano Ronaldo is a **soccer player**.
Consumption of alcohol is allowed only for **adults**.
A fraction consists of **numbers**.
Words consist of **letters**.
Tarantulas are **spiders**.
A boa constrictor is a **snake**.
Paper clips are **stationery**.
Ketchup is made from **tomatoes**.
A trunk consists of **wood**.
Steel is a **metal**.
Cheddar is a **cheese**.
A sofa can be found in the **living room**.
Lamps hang from the **ceiling**.
An hour has 60 **minutes**.
A light bulb works with **electricity**.
Sunscreen protects your **skin**.
Risotto is made from **rice**.
Fish live in **water**.
Eagles have **feathers**.
Shower gel makes you **clean**.
Umbrellas keep you **dry**.
Tankships are **heavy**.
A peach is a **fruit**.
Letters are written on **paper**.
You put a nail in the wall with a **hammer**.
Ice cubes are kept in the **freezer**.
A day starts in the **morning**.
You go to sleep **barefoot**.
People like to eat bread with **butter**.

APPENDIX B

Overview of true versions of sentences used in Experiment 2 and Experiment 3. False versions were created by switching the last word (highlighted) of Set A and Set B for each corresponding sentence.

Set A	Set B
East is where the sun rises.	**West** is where the sun sets.
A **slow** animal is a turtle.	**Fast** animals are cheetahs.
A **tall** animal is a giraffe.	**Short** describes the height of dwarfs.
Trees constitute a forest.	**Sand** is used by kids to build castles on the beach.
Cutting is done with scissors.	**Pasting** is done with glue.
One of the **countries** in the world is England.	An example for a **city** is Berlin.
Water is the element boats sail on.	**Earth** is the element horses run on.
In the **summer** people wear shorts.	In the **winter** people wear gloves.
Sweet is the taste of sugar.	**Bitter** is the taste of beer.
Sour is the taste of lemons.	**Salty** is the taste of potato chips.
Cold is the temperature in the Arctic.	**Hot** is the temperature of fire.
Scales cover the bodies of fish.	**Pores** can be found on the bodies of humans.
Feathers cover the bodies of birds.	**Fur** covers the bodies of dogs.
Liquid is the physical condition of wine.	**Solid** is the physical condition of walls.
Clean is what soap makes you.	**Dirty** is what grease makes you.
Dry is what you are after using a towel.	**Wet** is what you are after being in the rain.
Wheels are parts of a car.	**Legs** are parts of gazelles.
Bunnies hide Easter eggs.	**Chicken** are called roosters when they are male.
New York is the location of the Statue of Liberty.	**Paris** is the location of the Eiffeltower.
Peace was what John Lennon demonstrated for.	**War** reigned the world in 1939.
Green traffic lights signal you to cross the road.	**Red** is the color of blood.
Bright is a quality of flashlights.	It is **dark** when you switch off the light.
Big describes the size of elephants.	**Small** describes the size of ants.
The **sky** is where airplanes fly.	**Streets** are places where cars drive.
White is the color of snow.	**Black** is the color of a panther.
Hard is a feature of rocks.	**Soft** is a quality of silk.
Heavy describes the weight of jumbo jets.	**Light** describes the weight of feathers.
An example for a **fruit** is a pear.	An example for a **vegetable** is broccoli.
An example for a **dish** is baked beans.	An example for a **drink** is lemonade.
Expensive is the price of diamonds.	**Cheap** is what discounters are.
At **nigh**t time owls are active.	The **day** is the time when kids go to school.
Round describes the shape of peas.	**Square** describes the shape of a TV screen.
Twelve **months** constitute a year.	Twenty-four **hours** constitute a day.
Paper is the material magazines are made of.	**Textile** is the material clothes are made of.
A **key** is used to lock a door.	A **knife** is used to cut bread.

Silent is what the audience should be during a theatre performance.
Loud is what rock concerts are.

The **basement** is the lowest story of a building.
The **attic** can be reached by taking the stairs up.

An **oven** is used to bake cake.
A **fridge** is a place to keep ice cream.

Sickness is indicated by fever, coughing and a running nose.
Health means you can leave the hospital.

It is **rare** to win the lottery.
It is **common** to prefer holidays over workdays.

Tracks are the infrastructure trains run on.
Highways are the infrastructure buses run on.

Bees make honey.
Butterflies develop from caterpillars.

Right is the side of most people´s dominant hand.
Left is the side that people in Britain drive on.

Empty means you need to charge your battery.
Full is when the bath tub is overflowing.

The **south** is where Texas lies.
The **north** is where Alaska lies.

An example for **furniture** is a chair.
An example for **clothing** is a sweater.

An example for a **reptile** is an alligator.
An example for a **mammal** is a wolf.

Narrow describes the eye of a needle.
Wide describes the dimension of a baseball field.

Dollar is the currency in the USA.
Ruble is the currency in Russia.

Tails are wagged by dogs.
Whiskers are left and right from a cat's nose.

Happiness results from health, luck and enough money.
Sadness results from the death of a dear one.

In the **morning** people eat breakfast.
In the **evening** people eat dinner.

Spoons are used to eat soup.
Forks are used to eat salad.

The **door** is the place where you enter your house.
The **window** can be opened to let air inside the car while driving.

The **bedroom** is the place where people lie down to sleep.
The **kitchen** is the place where dinner is cooked.

Teeth are used to chew.
Your **tongue** is coated when you are ill.

A **cup** is used to drink coffee.
A **pan** is used to fry a steak.

The **neck** is the bodypart where scarfs are worn.
The **wrist** is the bodypart where watches are worn.

Feet are the bodyparts where shoes are worn.
Hands are the bodyparts where wedding rings are worn.

An example for an **actor** is Tom Hanks.
An example for a **singer** is Bob Marley.

Cans are used to sell tuna.
Bottles are used to sell water.

An example for a **number** is 4.
An example for a **letter** is Y.

Paws are body parts of tigers.
Hands are bodyparts of children.

Underground is where moles live.
Trees are where squirrels live.

Strings are parts of violins.
Keyboards are parts of accordions.

An example for a DIY **tool** is a scraper.
An example for a **musical instrument** is a harp.

An example for a **biped** is an ostrich.
An example for a **quadruped** is a horse.

An example for **nuts** are macadamia.
An example for **berries** are black currants.

Round is the shape of a tennis ball.
Square is the shape of a dice.

Milk is what a cow gives.
Juice is what you get when you squeeze oranges.

Halloween is when kids play 'trick or treat'.
Christmas is the time when Santa Claus brings presents.

Thorns are parts of roses.
Stings cover the bodies of hedgehogs.

Hives are places where bees live.
Caves are places where bears live.

A **grain** type is rye.
An **algae** type is seaweed.

Noses are used to smell.
Ears are used to hear.

In **Asia** you can find Japan.
In **Europe** you can find the Netherlands.

Vehicles are fixed in a garage.
Pottery is created with clay and a potter's wheel.

Democracy is the system of government in the USA.
Monarchy is the system of government in Saudi Arabia.

An example for a **painter** is Van Gogh.
A **whiskey** brand is Johnnie Walker.
A **wine** name is Cabernet Sauvignon.
Women can become pregnant.
Blue is the color of the sky.
The **moon** was visited by Neil Armstrong.
In **Spain** you can find Barcelona.
Watches are produced by Rolex.
A **zipper** is used to close a jacket.
The **trumpet** was played by Louis Armstrong.
A **broom** is used by a witch to fly.
A **pumpkin** turned into Cinderella's carriage.
An example for a **primate** is an orangutan.
A well-known **basketball** player is Kobe Bryant.
Children use sandboxes as playgrounds.
Letters constitute the alphabet.
Digits are used to denote one's birth date.
An example for a **snake** is a cobra.
An example for a **spider** is a black widow.
An example for **cosmetics** are lipsticks.
Potatoes are used to make chips.
Glass is used to make mirrors.
An example for **wood** is beech.
An example for **sausage** is pepperoni.
The **bathroom** is the place where a shower can be found.
Floors are swept with brooms.
Nine **months** is the duration of a pregnancy.
Fire is used to light candles.
Eyes can be protected by sunglasses.
Beans are what coffee is made from.
Earth is the element sheep live on.
Fur covers the bodies of hamsters.
Dirty is what mud makes you.
Wet is what you are after swimming.
Light describes the weight of balloons.
An example for a **vegetable** is a cucumber.
Textile is the material pants are made of.
A **saw** is used to cut wood.
A **microwave** is used to warm food.
The **evening** is the end of a day.
In **shoes** is how you leave the house.
Soap is used to wash your hands.

An example for a **poet** is Shakespeare.
A **soda** brand is Pepsi.
Vodka brands are Absolut and Smirnoff.
Men start having a beard at puberty.
Yellow is the color of corn.
Mars is the red planet.
In **Germany** you can find Hamburg.
An example for a **computer** is a Mac from Apple.
A **handle** is used to open a door.
The **guitar** was played by Jimi Hendrix.
A **carpet** is the magic ride of Aladdin.
Spinach is Popeye's food for strength.
An example for a **rodent** is a beaver.
A well-known **soccer** player is Cristiano Ronaldo.
Adults are allowed to consume alcohol.
Digits are part of a fraction.
Letters are part of a word.
An example for a **spider** is a tarantula.
An example for a **snake** is a boa constrictor.
An example for **stationery** are paper clips.
Tomatoes are used to make ketchup.
Wood is what a toothpick consists of.
An example for **metal** is steel.
An example for **cheese** is cheddar.
The **living room** is the place where a sofa can be found.
The **ceiling** is the place lamps hang from.
Sixty **minutes** constitute an hour.
Electricity is what makes a light bulb work.
Skin can be protected by sunscreen.
Rice is the main ingredient of risotto.
Water is the element fish live in.
Feathers cover the bodies of eagles.
Clean is what you are after using shower gel.
Dry is what you stay with an umbrella.
Heavy describes the weight of tankships.
An example for a **fruit** is a peach.
Paper is used to write letters.
Hammers are used to put nails in the wall.
The **freezer** is a place to keep ice cubes.
The **morning** is the beginning of a day.
Barefoot is how you go to sleep.
Butter is eaten with bread.

APPENDIX C

Overview of true versions of sentences used in Experiment 4. False versions were created by switching the last word (highlighted) of Set A and Set B for each corresponding sentence.

Set A	**Set B**
Bulls are male **cows**.	Roosters are male **chicken**.
Halloween is in **October.**	Christmas is in **December.**
Lemons are **sour**.	Chips are **salty**.
Macadamia are **nuts**.	Grapes are **berries**.
Chardonnay is a **wine**.	Smirnoff is a **vodka**.
You sleep in the **bedroom**.	You cook in the **kitchen**.
You chew with your **teeth**.	You lick with your **tongue**.
You cut with a **saw**.	You hit with a **hammer**.
You drink from a **cup**.	You fry in a **pan**.
Wasps are **insects**.	Pythons are **snakes**.
Fries are eaten with **ketchup**.	Bread is eaten with **butter**.
Trucks are **vehicles**.	Helicopters are **aircrafts**.
Socks are worn on **feet**.	Hats are worn on **heads**.
Federer plays **tennis**.	Clapton plays **guitar**.
Flour is kept in the **cupboard**.	Popsicles are kept in the **freezer**.
Picasso was a **painter**.	Shakespeare was a **poet**.
Cows give **milk**.	Chicken give **eggs**.
Bryant played **basketball**.	Maradona played **soccer**.
Manhattan is in **New York**.	Montmatre is in **Paris.**
Mirrors are **glass**.	Lumber is **wood**.
Chocolate contains **cocoa**.	Mayonnaise contains **eggs**.
Rolex produces **watches**.	Kellogg's produces **cereals**.
Earthworms live **underground**.	Jellyfish live **underwater**.
Iceland is a **democracy**.	Qatar is a **monarchy**.
Grass is **green**.	Blood is **red**.
Forests consist of **trees**.	Dunes consist of **sand**.
Kennedy was a **president**.	Elvis was a **singer**.
Earth is a **planet**.	Africa is a **continent**.
Japan is in **Asia**.	Denmark is in **Europe**.
Lipsticks are **cosmetics**.	Crayons are **stationery**.
Bacardi is a brand of **rum**.	Pepsi is a brand of **soda**.
Showers can be found in **bathrooms**.	Mattresses can be found in **beds**.
Jackets are opened with **zippers**.	Doors are opened with **handles**.
USA 's currency is **dollar**.	Russia's currency is **ruble**.
A candle is lit by **fire**.	A bulb is lit by **electricity**.
Honey comes from **bees**.	Wool comes from **sheep**.
Magazines are made from **paper**.	Clothes are made from **fabric**.
Pants are made from **fabric**.	Cartons are made from **paper**.
Screwdrivers are **tools**.	Guns are **weapons**.
Sugar is kept in the **cupboard**.	Yoghurt is kept in the **fridge**.
Scissors are used to **cut**.	Glue is used to **paste**.
England is a **country**.	Berlin is a **city**.

Candy is **sweet**.
Fish breathe through the **gills**.
Birds have **feathers**.
Cars have **wheels**.
Veins transport **blood**.
A year ends with **December**.
Scissors cut **hair**.
Trains run on **tracks**.
A chair is a piece of **furniture**.
Alligators are **reptiles**.
Soup is eaten with a **spoon**.
Scarves are worn around the **neck**.
Tuna is sold in **cans**.
Tigers have **paws**.
Violins have **strings**.
Ostriches are **bipeds**.
Roses have **thorns**.
Bees live in **hives**.
Rye is a **grain**.
You smell with your **nose**.
The sky is **blue**.
Barcelona is in **Spain**.
Orangutans are **primates**.
Cobras are **snakes**.
Chips are made from **potatoes**.
Oak is a **wood**.
Pregnancy lasts **months**.
Sunglasses protect your **eyes**.
Coffee is made from **beans**.
Hamsters have **fur**.
The sun is **shining**.
Turtles are **slow**.
Giraffes are **tall**.
Clouds are made of **water**.
Swimsuits are worn in **water**.
Gloves keep you **warm**.
Silk is **soft**.
Soap makes you **clean**.
Towels make you **dry**.
A dove is a sign of **peace**.
Flashlights are **bright**.
Elephants are **massive**.
Snow is **white**.
Rocks are **hard**.
Tankships are **heavy**.
A pear is a **fruit**.
Pizza is a **dish**.
Pigs live in **barns**.

Beer is **bitter**.
Humans breathe through the **nose**.
Dogs have **fur**.
Cats have **legs**.
Bees transport **pollen**.
A week ends with **Sunday**.
Knives cut **bread**.
Buses run on **highways**.
A sweater is a piece of **clothing**.
Wolves are **mammals**.
Salad is eaten with a **fork**.
Watches are worn around the **wrist**.
Water is sold in **bottles**.
Children have **hands**.
Accordions have **keyboards**.
Horses are **quadrupeds**.
Hedgehogs have **stings**.
Bears live in **caves**.
Seaweed is an **alga**.
You hear with your **ears**.
Corn is **yellow**.
Hamburg is in **Germany**.
Beavers are **rodents**.
Tarantulas are **spiders**.
Ketchup is made from **tomatoes**.
Steel is a **metal**.
Puberty lasts **years**.
Sunscreen protects your **skin**.
Risotto is made from **rice**.
Eagles have **feathers**.
The wind is **blowing**.
Fireworks are **loud**.
Ants are **small.**
Roads are made of **asphalt**.
Pajamas are worn in **bed**.
Umbrellas keep you **dry**.
Walls are **firm**.
Food makes you **full**.
Grease makes you **dirty**.
A skull is a sign of **poison**.
Candlelight is **romantic.**
Feathers are **light**.
Strawberries are **red**.
Water is **liquid**.
Mice are **little**.
Cheddar is a **cheese**.
Lemonade is a **drink**.
Snails live in **shells**.

Humans eat **doughnuts**.
Peas are **round**.
Moles are **blind**.
Baking requires an **oven**.
Exercise makes you **fit**.
Flowers are a common **present**.
Nile is a **river**.
Void means **empty**.
Comedy is a genre of **movies**.
Sahara is a **desert**.
Smiling is a sign of **happiness**.
The week starts with **Monday**.
A house has a **door**.
Shoes are worn on **feet**.
Four is a **number**.
Coins are **flat**.
Embryos grow in **wombs**.
You clean yourself with **shower gel**.
The alphabet consists of **letters**.
The Bible is a famous **book**.
Peperoni is a **sausage**.
Kindness is a **virtue**.
Birds fly in the **sky**.
Mud makes you **dirty**.
Rain is **wet**.
Fleas are **tiny**.
A cucumber is a **vegetable**.
Blue is a **color**.
Apples hang on **trees**.
Pumpkins are **orange**.

Goats eat **grass**.
Pyramids are **triangular**.
Crocodiles are **dangerous**.
Driving requires a **license**.
A virus makes you **sick**.
The flu is a common **disease**.
Everest is a **mountain**.
Overcrowded means **replete**.
Rock is a genre of **music**.
Atlantic is an **ocean**.
Shivering is a sign of **fear**.
The year starts with **January**.
A face has a **mouth**.
Eyeglasses are worn on the **nose**.
Sour is a **flavor**.
Dice are **square**.
Plants grow in **pots**.
You warm yourself with a **blanket.**
A week consists of **days**.
The Kiss is a famous **painting**.
Broccoli is a **vegetable**.
Smell is a **sense.**
Fish live in **water**.
Fasting makes you **hungry**.
Fire is **hot**.
Tankships are **heavy**.
Cream is **dairy**.
Ten is a **number**.
Washing hangs on **clotheslines.**
Spinach is **green**.

References

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior*, *22*(3), 261-295. doi.org/10.1016/S0022-5371(83)90201-3

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390-412. doi.org/10.1016/j.jml.2007.12.005

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. doi.org/10.1016/j.jml.2012.11.001

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. hdl.handle.net/10.18637/jss.v067.i01

Berkum, J. J. V., Hagoort, P., & Brown, C. M. (1999). Semantic integration in sentences and discourse: Evidence from the N400. *Journal of Cognitive Neuroscience*, *11*(6), 657-671. doi.org/10.1162/089892999563724

Biderman, D., & Mudrik, L. (2017). Context modulation of ambiguous object perception in the absence of awareness. *Journal of Vision*, *17*(10), 1224-1224. doi.org/10.1167/17.10.1224

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

Brandman, T., & Peelen, M. V. (2017). Interaction between scene and object processing revealed by human fMRI and MEG decoding. *Journal of Neuroscience*, 0582-17. doi.org/10.1523/JNEUROSCI.0582-17.2017

Chan, M. P. S., Jones, C. R., Hall Jamieson, K., & Albarracín, D. (2017). Debunking: A meta-analysis of the psychological efficacy of messages countering misinformation. *Psychological Science, 28,* 1531-1546. doi.org/10.1177/0956797617714579

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, *12*(4), 335-359. doi.org/10.1016/S0022-5371(73)80014-3

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological review*, *82*(6), 407-428. doi/10.1037/0033-295X.82.6.407

Connell, L., & Keane, M. T. (2006). A model of plausibility. *Cognitive Science*, *30*(1), 95-120. doi.org/10.1207/s15516709cog0000_53

Conroy, N. J., Rubin, V. L., & Chen, Y. (2015). Automatic deception detection: Methods for finding fake news. *Proceedings of the Association for Information Science and Technology*, *52*(1), 1-4. doi:10.1002/pra2.2015.145052010082

Cook, A. E., Walsh, E. K., Bills, M. A., Kircher, J. C., & O'Brien, E. J. (2016). Validation of semantic illusions independent of anomaly detection: Evidence from eye movements. *The Quarterly Journal of Experimental Psychology*, 1-11. doi.org/10.1080/17470218.2016.1264432

Cook, A. E., & O'Brien, E. J. (2014). Knowledge activation, integration, and validation during narrative text comprehension. *Discourse Processes*, *51*(1-2), 26-49. doi.org/10.1080/0163853X.2013.855107

Davenport, J. L., & Potter, M. C. (2004). Scene consistency in object and background perception. *Psychological Science, 15*(8), 559-564. doi.org/10.1111%2Fj.0956-7976.2004.00719.x

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience, 8*(8), 1117-1121. doi:10.1038/nn1504

Erickson, T. D., & Mattson, M. E. (1981). From words to meaning: A semantic illusion. *Journal of Verbal Learning and Verbal Behavior*, *20*(5), 540-551. doi.org/10.1016/S0022-5371(81)90165-1

Eslick, A. N., Fazio, L. K., & Marsh, E. J. (2011). Ironic effects of drawing attention to story errors. *Memory*, 19, 184–191. doi.org/10.1080/09658211.2010.543908

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,* 175-191. doi.org/10.3758/BF03193146

Fazio, L. K., Barber, S. J., Rajaram, S., Ornstein, P. A., & Marsh, E. J. (2013). Creating illusions of knowledge: Learning errors that contradict prior knowledge. *Journal of Experimental Psychology: General, 142,* 1–5. dx.doi.org/10.1037/a0028649

Ferreira, F., Bailey, K. G., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science*, *11*(1), 11-15. doi.org/10.1111/1467-8721.00158

Gilbert, D. T. (1991). How mental systems believe. *American Psychologist*, *46*(2), 107-119. dx.doi.org/10.1037/0003-066X.46.2.107

Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science*, *304*(5669), 438-441. doi:10.1126/science.1095455

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

Helfand, D. J. (2016). *A survival guide to the misinformation age: Scientific habits of mind*. Columbia University Press.

Hinze, S. R., Slaten, D. G., Horton, W. S., Jenkins, R., & Rapp, D. N. (2014). Pilgrims sailing

the Titanic: Plausibility effects on memory for misinformation. *Memory & Cognition, 42*(2), 305-324. doi.org/10.3758/s13421-013-0359-9

Hutchison, K. A. (2003). Is semantic priming due to association strength or feature overlap? A microanalytic review. *Psychonomic Bulletin & Review, 10*(4), 785-813. doi.org/10.3758/BF03196544

Inquisit 5.0.11.0 [Computer software]. (2018). Seattle, WA: Millisecond Software.

Isberner, M. B., & Richter, T. (2013). Can readers ignore implausibility? Evidence for nonstrategic monitoring of event-based plausibility in language comprehension. *Acta Psychologica, 142*(1), 15-22. doi.org/10.1016/j.actpsy.2012.10.003

Isberner, M. B., & Richter, T. (2014). Does validation during language comprehension depend on an evaluative mindset? *Discourse Processes, 51*(1-2), 7-25. doi.org/10.1080/0163853X.2013.855867

Jacoby, L. L., & Kelley, C. M. (1987). Unconscious influences of memory for a prior event. *Personality and Social Psychology Bulletin, 13*(3), 314-336. doi.org/10.1177/0146167287133003

Jeffreys, H. (1961). *Theory of probability* (3rd ed.). Oxford: Oxford University Press.

JASP Team (2018). JASP (Version 0.9) [Computer software].

Kendeou, P. (2014). Validation and comprehension: An integrated overview. *Discourse Processes, 51*(1-2), 189-200. doi.org/10.1080/0163853X.2013.855874

Kirkham, R. L. (1992). *Theories of Truth*. Cambridge, MA: MIT Press.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207*(4427), 203-205. doi:10.1126/science.7350657

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature, 307*(5947), 161-163. doi.org/10.1038/307161a0

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the

N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, *62*, 621-647. doi.org/10.1146/annurev.psych.093008.131123

Lazer, D. M., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., Metzger, M. J., Nyhan, B., Pennycook, G., Rothschild, D., Schudson, M., Sloman, S. A., Sunstein, C. R., Thorson, E. A., Watts, D. J., & Zittrain, J. L. (2018). The science of fake news. *Science*, *359*(6380), 1094-1096. doi.org/10.1126/science.aao2998

Lee, M. D., & Wagenmakers, E.-J. (2013). Bayesian cognitive modeling: A practical course. New York, NY, US: Cambridge University Press. dx.doi.org/10.1017/CBO9781139087759

Lenth, R. (2016). Least-Squares Means: The R Package lsmeans. *Journal of Statistical Software, 69*(1), 1 - 33. dx.doi.org/10.18637/jss.v069.i01

Lewandowsky, S., Ecker, U. K., 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. doi.org/10.1177/1529100612451018

Loersch, C., & Payne, B. K. (2011). The situated inference model an integrative account of the effects of primes on perception, behavior, and motivation. *Perspectives on Psychological Science, 6,* 234-252. doi.org/10.1177/1745691611406921

Marsh, E. J., & Fazio, L. K. (2006). Learning errors from fiction: Difficulties in reducing reliance on fictional stories. *Memory & Cognition*, *34*, 1140–1149. doi:10.3758/BF03193260

Marsh, E. J., Meade, M. L., & Roediger, H. L. (2003). Learning facts from fiction. *Journal of Memory and Language, 49*, 519–536. doi.org/10.1016/S0749-596X(03)00092-5

Masson, M. E. (1995). A distributed memory model of semantic priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(1), 3-23. doi.org/10.1037/0278-7393.21.1.3

Mayo, R. (2015). Cognition is a matter of trust: Distrust tunes cognitive processes. *European Review of Social Psychology, 26*(1), 283–327. doi.org/10.1080/10463283.2015.1117249

McNamara, T. P. (2005). *Semantic priming: Perspectives from memory and word recognition*. New York: Psychology Press doi.org/10.4324/9780203338001

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90*(2), 227-234. doi.org/10.1037/h0031564

Moss, H. E., Hare, M. L., Day, P., & Tyler, L. K. (1994). A distributed memory model of the associative boost in semantic priming. *Connection Science*, *6*(4), 413-427. doi.org/10.1080/09540099408915732

Neely, J. 1977. Semantic priming and retrieval from lexical memory: Role of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General, 106,* 226-254. doi.org/10.1037//0096-3445.106.3.226

Neely, J. H. (2012). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In *Basic Processes in Reading* (pp. 272-344). Routledge. doi.org/10.4324/9780203052242

Newman, E. J., Garry, M., Unkelbach, C., Bernstein, D. M., Lindsay, D. S., & Nash, R. A. (2015). Truthiness and falsiness of trivia claims depend on judgmental contexts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*(5), 1337-1348. doi.org/10.1037/xlm0000099

O'Brien, E. J., & Cook, A. E. (2016). Coherence threshold and the continuity of processing: The RI-Val model of comprehension. *Discourse Processes, 53*(5-6), 326-338. doi.org/10.1080/0163853x.2015.1123341

Orenes, I., & Santamaria, C. (2014). Visual content of words delays negation. *Acta Psychologica, 153,* 107-112. doi.org/10.1016/j.actpsy.2014.09.013

Palan, S., & Schitter, C. (2018). Prolific. ac—a subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, *17*, 22-27. doi.org/10.1016/j.jbef.2017.12.004

Pantazi, M., Kissine, M., & Klein, O. (2018). The power of the truth bias: False information affects memory and judgment even in the absence of distraction. *Social Cognition*, *36*(2), 167-198. doi.org/10.1521/soco.2018.36.2.167

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

Piai, V., Roelofs, A., & van der Meij, R. (2012). Event-related potentials and oscillatory brain responses associated with semantic and Stroop-like interference effects in overt naming. *Brain Research*, *1450*, 87-101. doi.org/10.1016/j.brainres.2012.02.050

Posner, M., & Snyder, C. R. R. (1975). Facilitation and inhibition in the processing of signals. In P. M. A. Rabbitt and S. Dornic (Eds.), *Attention and performance V*. New York: Academic Press. doi.org/10.2307/1421416

Pulvermüller, F. (2013). How neurons make meaning: brain mechanisms for embodied and abstract-symbolic semantics. *Trends in Cognitive Sciences*, *17*(9), 458-470. doi.org/10.1016/j.tics.2013.06.004

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.r-project.org/

Rapp, D. N. (2016). The consequences of reading inaccurate information. *Current Directions in Psychological Science*, *25*(4), 281-285. doi.org/10.1177/0963721416649347

Rapp, D. N., & Braasch, J. L. (2014). *Processing inaccurate information: Theoretical and applied perspectives from cognitive science and the educational sciences*. Cambridge, MA: MIT Press.

Rapp, D. N., Hinze, S. R., Kohlhepp, K., & Ryskin, R. A. (2014). Reducing reliance on inaccurate information. Memory & Cognition, 42(1), 11-26. https://doi.org/10.3758/s13421-013-0339-0

Rapp, D. N., Hinze, S. R., Slaten, D. G., & Horton, W. S. (2014). Amazing stories: Acquiring and avoiding inaccurate information from fiction. *Discourse Processes*, *51*(1-2), 50–74. doi.org/10.1080/0163853X.2013.855048

Reber, R., & Schwarz, N. (1999). Effects of perceptual fluency on judgments of truth. *Consciousness and Cognition*, *8*(3), 338-342. doi.org/10.1006/ccog.1999.0386

Richter, T. (2015). Validation and comprehension of text information: Two sides of the same coin. *Discourse Processes*, *52*(5-6), 337-355. doi.org/10.1080/0163853X.2015.1025665

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

Richter, T., Schroeder, S., & Wöhrmann, B. (2009). You don't have to believe everything you read: Background knowledge permits fast and efficient validation of information. *Journal of Personality and Social Psychology*, *96*(3), 538-558. dx.doi.org/10.1037/a0014038

Sanford, A. J. (2002). Context, attention and depth of processing during interpretation. *Mind and Language*, *17*(1-2), 188-206. doi.org/10.1111/1468-0017.00195

Schul, Y., Mayo, R., & Burnstein, E. (2004). Encoding under trust and distrust: The spontaneous activation of incongruent cognitions. *Journal of Personality and Social Psychology, 86*(5), 668–679. dx.doi.org/10.1037/0022-3514.86.5.668

Singer, M. (2013). Validation in reading comprehension. *Current Directions in Psychological Science*, *22*(5), 361-366. doi.org/10.1177/0963721413495236

Singer, M. (2019). Challenges in Processes of Validation and Comprehension. *Discourse Processes*, 1-19. https://doi.org/10.1080/0163853X.2019.1598167

Singer, M., & Doering, J. C. (2014). Exploring individual differences in language validation. *Discourse Processes*, *51*(1-2), 167-188. doi.org/10.1080/0163853X.2013.855534

Traxler, M. J., Foss, D. J., Seely, R. E., Kaup, B., & Morris, R. K. (2000). Priming in sentence processing: Intralexical spreading activation, schemas, and situation models. *Journal of Psycholinguistic Research*, *29*(6), 581-595. doi.org/10.1023/A:1026416225168

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. doi.org/10.1037/0278-7393.33.1.219

[dataset] Weil, R. & Mudrik, L. (2018). Detecting falsehood relies on mismatch detection between sentence components. Open Science Framework (OSF). DOI 10.17605/OSF.IO/C6J4B

Weil, R., Schul, Y. & Mayo, R. (2019). Correction of evident falsehood requires explicit negation. *Journal of Experimental Psychology: General.* Advance online publication. http://dx.doi.org/10.1037/xge0000635

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology, 9*, 1-27. doi.org/10.1037/h0025848