The Epistemic Status of Processing Fluency as Source for Judgments of Truth

Rolf Reber · Christian Unkelbach

Published online: 7 September 2010
© The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract This article combines findings from cognitive psychology on the role of processing fluency in truth judgments with epistemological theory on justification of belief. We first review evidence that repeated exposure to a statement increases the subjective ease with which that statement is processed. This increased processing fluency, in turn, increases the probability that the statement is judged to be true. The basic question discussed here is whether the use of processing fluency as a cue to truth is epistemically justified. In the present analysis, based on Bayes’ Theorem, we adopt the reliable-process account of justification presented by Goldman (1986) and show that fluency is a reliable cue to truth, under the assumption that the majority of statements one has been exposed to are true. In the final section, we broaden the scope of this analysis and discuss how processing fluency as a potentially universal cue to judged truth may contribute to cultural differences in commonsense beliefs.

It was Napoleon, I believe, who said that there is only one figure in rhetoric of serious importance, namely, repetition. The thing affirmed comes by repetition to fix itself in the mind in such a way that it is accepted in the end as a demonstrated truth.

Gustave Le Bon (1895/1960, p. 125)
1 Introduction

When people hear or see a statement repeatedly, they believe that this statement is more likely to be true than new statements which they have never encountered before. As the quote from Le Bon (1895) indicates, this fact has been known for a long time. A classical empirical demonstration of this effect was given by Hasher et al. (1977). In a first session, participants heard forty statements, half of which were repeated later. Some statements were true, for example, “Lithium is the lightest of all metals”, some of them were false, for example, “The people’s Republic of China was founded in 1947”. Participants judged each statement’s truth on a scale from 1 (definitely false) to 7 (definitely true). In a second session, participants heard forty statements again. Twenty of them were presented in the first session and twenty were new statements, and participants again judged their truth. Statements heard twice were judged as being more probably true than new statements. This robust finding supports Le Bon’s claim and has been replicated many times by different laboratories (e.g., Arkes et al. 1989; Bacon 1979; Begg et al. 1992; Brown and Nix 1996; see Dechêne et al. 2009b, for a review).

There are two routes from repeated exposure to truth judgments: First, people might recall that they have heard a statement before and conclude that this statement must be true. This route to judge a statement as “true” is based on convergent validity (Arkes et al. 1991). Indeed, Brown and Nix (1996), who used the same basic paradigm as Hasher et al. (1977), could show that after a long delay between presentation of statements—that were said to be true—and truth judgment, participants judged a statement to be more probably true if they could recognize that it had been presented before. Second, however, such repeated statements were also judged as being more probably true if participants were not able to recognize the statement and thus did not refer their truth judgments back to the statement in the first session. This non-referential part of the repetition-truth link will interest us in the remainder of the article. How does this link work?

If people encounter a statement repeatedly, they can process it more efficiently; the statement is processed more fluently (e.g., Jacoby and Dallas 1981). Processing fluency is defined as the subjective experience of ease with which a stimulus is processed. If a person cannot recognize the statement, this experienced ease is taken as information when judging the truth of a statement. If the statement can be processed easily, the person will conclude that the statement is true; if the statement is difficult to process, she concludes that the statement is not true. This non-referential effect of repetition on judgments of truth is a well-established effect (e.g., Begg et al. 1992; Reber and Schwarz 1999; Unkelbach 2007; Unkelbach and Stahl 2009).

The basic question we address in this article is whether processing fluency is an epistemically reliable¹ cue to truth. In other words, under which conditions is it a good strategy to judge truth using processing fluency as a cue? In order to discuss

---

¹ Goldman (1986) noted that reliability as used in the literature on epistemology corresponds to the term validity in the behavioral sciences. We use these two expressions interchangeably, but consistently insofar as we use reliability when we refer to epistemological literature, and validity when we refer to psychological literature.
this question, we shall first describe evidence for both determinants of fluency and the fluency-truth-link before we proceed to discuss how the epistemic justification of fluency can be evaluated. From the different philosophical theories of truth (see Kirkham 1992) that could be used to evaluate the reliability of fluency as a cue to truth, we choose the reliable-process approach by Alvin Goldman (1986). We describe psychological studies that are pertinent to epistemic justification of fluency and discuss how the reliable-process theory could contribute to our main question. In a final outlook, we broaden the scope and discuss how research on the link of fluency to judged truth in factual statements may inform the metacognitive basis of public opinions and doxastic beliefs.

1.1 Determinants of Fluency

There are three determinants of processing fluency that are well established in the cognitive literature: repetition, structure, and perceptual features. In the following, we give examples of how these factors influence processing fluency.

Repetition Evidence abounds that repeated exposure to stimuli results in a long-lasting increase in fluent processing of these stimuli. Repeated exposure facilitates processing of the stimuli in subsequent tasks; for example, masked words are more likely to be identified if they were previously presented on a study list (Jacoby and Dallas 1981). These repetition effects are largely independent from manipulations that typically affect conscious retrieval from memory. For example, elaborating on a stimulus increases conscious retrieval, but does not affect measures of non-referential measures of memory (Begg et al. 1985). Other facilitating effects of repetition were observed using a clarification procedure. In this procedure, test words slowly emerge from a pattern mask, becoming more and more readable. In one study, participants’ task was to press a button as soon as they could identify the word (normally after a couple of seconds) and then to write it down. Accurate identification was faster for repeated words than non-repeated words (Feustel et al. 1983). Moreover, repeated stimuli elicit less attentional orienting than novel stimuli (Desimone et al. 1995), further bolstering the notion that stimulus repetition increases processing fluency.

Implicit learning of structure After A. S. Reber (1967) presented his study about implicit learning of an artificial grammar, a debate emerged over the mechanism behind the observed phenomena. One of the mechanisms proposed was processing fluency. In the study by Reber (1967), participants were exposed to letter strings that are exemplars of a finite state grammar, and then were presented with novel letter strings; half of them followed the same grammar, the other half were not grammatical. Participants were able to classify the items into grammatical and non-grammatical stimuli at above-chance level. Indeed, research shows that grammatical stimuli are processed with higher fluency. For example, Buchner (1994) first exposed participants to grammatical letter strings and subsequently presented strings with the perceptual clarification procedure by Feustel et al. (1983), asking participants to identify the strings. Specifically, the test strings were presented in a black mask, from which pixels were gradually removed at random until the
participant could identify the letter string. Participants reacted faster to grammatical rather than ungrammatical letter strings, showing that grammatical letter strings were easier to process.

The fact that implicit learning of structure increases processing fluency generalizes the findings on exact repetition of a stimulus to similarity of structure. Even if people read a statement which is similar, but not exactly the same as statements read before, they may feel increased fluency and conclude that the statement is true. Again, processing fluency is a valid cue to structural similarity. However, in this case, it is difficult to determine whether structural similarity indeed is connected to truth: even if the statements one has encountered originally have been true, it is easily conceivable that a structurally similar statement does not correspond with reality. This may be illustrated by the study of Deese (1959; Roediger and McDermott 1995) who observed that participants claim to have been presented with a word from a list of similar words even if they have not seen it before. For example, people claimed to have been presented with the word sleep if they were presented with a word list of associates of the word sleep, such as bed, dream, and snooze. Similarly, as people can assess global fluency, but not disentangle its different sources, they may think that they have read the same statement before although fluency stemmed from reading a similar statement. However, there is a lack of findings on similarity and judged truth so that this issue awaits further research.

Perceptual and linguistic features

The first two determinants—repetition and implicit learning of structure—were based on earlier experiences of a person. In the perceptual world, there are many features that increase perceptual fluency, which is the subjective ease with which a stimulus can be perceived. Such features include figure-ground contrast, symmetry, or informational simplicity.

When running studies with linguistic stimuli, rhyming (McGlone and Tofighbakhsh 2000), text coherence (e.g., Rawson and Dunlosky 2002), orthographic regularity (Whittlesea and Williams 1998), or expectation that an expression follows a sentence stem, such as TRAIL in “The hikers got lost when they lost the TRAIL” versus “He helped the company build a new TRAIL” (Whittlesea 1993; Whittlesea and Williams 2001) has been used to manipulate processing fluency. Interestingly, people seem to be unable to dissociate processing fluency stemming from different sources, for example, previous experiences and perceptual features. Manipulating any of the described variables has basically the same effect: Fluently processed statements are judged to be true with higher probability.

1.2 Fluency and Truth: Empirical Evidence

The first to directly test the fluency-truth link were Begg et al. (1992), who presented half of the statements in the first session from a male and half from a female source. They informed participants that statements from a male source were true and statements from a female source were false, or vice versa. In the subsequent session, however, old statements labeled false were significantly more likely to be judged true than new statements. This means that participants were unable to retrieve the source of the statements, but still judged these statements true at above-baseline
level. Begg and colleagues (1992) concluded that participants used the experience elicited by the statements, that is, processing fluency, to judge their truth.

Other studies have addressed the fluency-truth link more directly: Reber and Schwarz (1999) argued that if repeated exposure to a statement increases its processing fluency and if people use processing fluency as a cue to judge a statement’s truth, then any fluency manipulation should increase the probability to judge a statement as “true”. In order to test this assumption, they presented statements of the form “Osorno is in Chile”, one by one, on a computer screen. The statements had strong or moderate contrast to the background, and were thus more or less fluently processed. Participants had to decide whether the statements were true or false. As predicted from a processing fluency account, statements were more likely to be judged “true” when presented in strong rather than moderate contrast. Similarly, McGlone and Tofighbakhsh (2000) showed that the same aphorisms were more likely to be perceived as true when presented in rhyming (e.g., “Birds of a feather flock together”) rather than nonrhyming form (e.g., “Birds of a feather flock conjointly”). They argued that rhyming also enhances the fluency with which statements are understood (see Rubin 1995), and fluency due to rhyming is used as a cue to judge the truth of the aphorism.

Expectations and attributions moderate the effects of fluency on judgments. Whittlesea and Williams (1998), for example, used words, regular nonwords, and irregular nonwords as stimuli. Regular nonwords had no meaning in English, but were nonetheless easy to read, such as HENSION or WIMBER. Irregular nonwords were difficult to read, such as STOFWUS or OUETIS. Participants first read words, regular nonwords, and irregular nonwords in a training phase. In a subsequent test phase, they were shown old and new stimuli from each of the three categories and had to perform three tasks on each stimulus: First, they had to pronounce it, then decide if it was a word or a nonword, and finally to decide if it had been presented in training. Processing fluency, as measured by pronunciation latency and lexical decision, was higher for words than for regular nonwords and higher for regular nonwords than for irregular nonwords. If processing fluency alone determined the proportion of false alarms in the recognition test, words would yield more false alarms than regular nonwords, which, in turn, would yield more false alarms than irregular nonwords. However, regular nonwords produced more false alarms than did words and irregular nonwords, respectively. Whittlesea and Williams (1998) concluded that not processing fluency per se, but the discrepancy between expected and actual fluency resulted in an increase in positive recognition judgments. It is not surprising that words can be processed fluently, whereas irregular nonwords cannot. However, fluent processing of regular nonwords comes as a surprise because there is a discrepancy between expected and actual fluency. This unexplained fluency is interpreted as familiarity, increasing the proportion of false alarms. In a similar vein, Hansen et al. (2008; see also Dechêne et al. 2009a), found that effects of fluency on judged truth were especially strong when statements were processed with surprising ease, that is, if high fluency statements were presented after a sequence of moderately fluent statements.

Why does ease of processing result in increased belief in a statement? Reber et al. (2004) assumed that processing fluency is hedonically marked and therefore always positive. In a similar vein, and in line with the theory of relevance proposed
by Sperber and Wilson (1986), one could argue that processing fluency became epistemically marked in the course of human evolution, which means that fluently processed statements are judged as being true. As cognitive systems aim at extracting information with the lowest effort possible, high fluency—or lack of effort—contributes to the relevance of a statement (Sperber and Wilson 1986).

However, several studies have shown that the link between fluency and truth is not fixed. In studies reported by Skurnik et al. (2000), the effect of fluency on truth was mediated by naïve beliefs: When participants believed that a majority of familiar statements was true, they judged familiar statements as being true; if they believed that a majority of familiar statements were false, they used familiarity as a cue to falsehood (see Winkielman and Schwarz 2001; Unkelbach and Stahl 2009, for related findings).

Using a more indirect manipulation, Unkelbach (2007) showed that people can learn to use processing fluency as a cue for truth but also for falseness. In a training phase, participants saw statements of which the factual truth status (true/false) was known. In addition, the fluency of the statements was manipulated by color contrast. In one condition, all fluently processed statements were true in the training phase (e.g., “A day has 24 h.”), and all non-fluently processed statements were false (e.g., “Aluminum is heavier than gold.”). At a subsequent test phase, statements of unclear factual status (e.g., “Cactuses can procreate via parthogenesis.”) were presented in high and low contrast and the standard pattern emerged: Easy to process statements were judged to be true with higher probability than difficult to process statements. However, in another condition, all fluently processed statements were false in the training phase and all non-fluently processed statements were true. This led to a reversal at test: Difficult to process statements were judged to be true with higher probability than easy to process statements. Furthermore, the reversal even emerged when fluency was manipulated by contrast in the training phase and repetition in the test phase. In other words, when participants learned that fluency due to contrast is related to falseness, they interpreted fluency at test which stemmed from a different source, namely repetition, in terms of falseness. Based on these results, the author argued that people usually judge fluently processed statements as true because people learn that there is a positive correlation between fluency and truth in real-world environments. Thus, fluency is an ecologically valid cue to truth if one assumes that more of the statements one encounters are true rather than untrue. This leads us directly to the issue of epistemic justification.

2 Justification of Beliefs

In order to define justification, we first have to ask: justification as what (see Kirkham 1992). A belief can be justified as being true. We then speak of epistemic justification which will be the kind of justification that interests us most. There are

---

2 One might argue that the malleability of the fluency experience undermines the claim that it is an ecologically valid cue; however, participants probably learned the correlation of the experience with a criterion within the context of the experiment, and the context triggers the interpretation of the experience. For illustration, when reading a murder mystery, experienced and learned readers know that the easy and obvious solution is never true, just because in the context of the murder mystery, the murderer is never the butler (see also the discussion in Unkelbach, 2007).
two other main kinds of justification, and it is important not to mix them with epistemic justification. The first is moral justification. We may find that a belief is morally justified, regardless of whether or not it is true, and we may think that a belief is true, but not morally justified. Another kind of non-epistemic justification is the utilitarian notion that a belief is justified if it has maximal expected utility. William James (1909) put forward a utilitarian theory of justification, but it is relatively easy to refute it as a theory of epistemic justification (see also Kirkham 1992): Beliefs may be true, even if they are damaging. For example, if a woman suffers from a risk of heart disease, she may profit from thinking that she has no risk of heart disease, because the belief that she has this risk may increase her blood pressure and thus contribute to the very risk she believes in. Although not believing in the risk may maximize one’s utility, it is not justified as being true; this shows that even if it is justified to have a belief on utilitarian grounds, it may not be justified as being true. The perhaps most well-known example is the so-called “Pascal’s wager” (Pascal 1678/1941). If we know the gains and losses that the belief in a statement brings, we may infer how adaptive a belief is for its bearer, but good adaption does not necessarily mean epistemic justification.

2.1 Truth

We characterize the concept of truth as a correspondence between a belief and a state of affairs (Kirkham 1992, pp. 119–130). For our purpose, we are going to assume that the correspondence is best covered by the notion of correlation, a notion that lies in the nature of a reliable-process theory of truth. We follow the line of Goldman (1986), who relied on Quine’s (1969) postulate of naturalized epistemology and proposed a reliable-process account in that a belief is justified if it comes from cognitive processes that yield true over false answers at an above-chance rate. According to Goldman (1986), a justificational rule system R is right if and only if R permits certain (basic) psychological processes, and the instantiation of these processes would result in a truth ratio of beliefs that meets some specified high threshold.\(^3\) We set this threshold at above-chance level: If people evaluate statements as being true or not true, with a chance level of 50% of being accurate, processing fluency is a reliable cue to truth if it results in significantly more than 50% correct decisions. This is a realist position insofar as the belief is related to a state of affairs that is independent of the belief. Therefore, it is not a coherentist theory of truth. However, Goldman argued that although coherence is not a required criterion for the justification of beliefs, it is a derivative standard. If a belief is true, it is required that it is consistent with other true beliefs. This is why a reliable-process theory of truth implies the requirement of coherence of beliefs.

2.2 Proximal Cues and Distal Criteria

The correspondence between a belief and a state of affairs can be broken down into two separate correlations: The correlation between a belief to the distribution of

\(^3\) Goldman (1986) distinguishes primary justification from secondary justification. The former refers to processes, the latter to methods that lead to accurate results. We are dealing with mental processes and therefore with primary justification.
proximal cues, and between proximal cue and a distal criterion (see Brunswik 1955, 1957). The distal criterion is a property of the environment, for example, the frequency of stomach cancer as a cause of death. However, the distal criterion is often not directly accessible; we have no sensory organs that tell us about this frequency. Thus, this representation is mediated by proximal (i.e., accessible) cues, which often results in a biased representation of the distal criterion in the human mind. The difference between the proximal cue-mediated representation of the criterion and the factual criterion is illustrated by a study of Lichtenstein et al. (1978). Participants were asked to judge the frequency of homicide and stomach cancer as causes of death. In the external world, stomach cancer is about five times more frequent than homicide; this is the distal criterion. However, people are more often exposed to instances of homicide than instances of stomach cancer, because media coverage for homicide is much higher than for stomach cancer; thus, the proximal cues do not convey a representative picture, and their distribution differs from the actual distal criterion. As a consequence, there are two separate sources of judgmental bias (see Fiedler 2000): First, the representation derived from the proximal cues differs from the actual distal criterion, just because the cues do not represent the environment reliably, as shown in the example above. Second, the representation delivered by the proximal cues may result in a biased representation, because processing of the cues is biased. For example, reports on homicides are often more vivid than reports on stomach cancer. Hence, even if the proximal frequency of news reports on stomach cancer and homicide corresponds to the actual distribution, that is, the distal criterion, people may be biased towards the frequency of homicide because these are represented more vividly than instances of stomach cancer. Beyond Lichtenstein et al. (1978), only few researchers have taken ecological data about the distribution of distal criterion into account. A study by Goldstein and Gigerenzer (2002) illustrates how this approach is relevant to epistemic justification.

2.3 The Recognition Heuristic

Let us illustrate epistemic justification with the recognition heuristic (Goldstein and Gigerenzer 2002). If asked which city has a larger population, San Antonio or San Diego, American students, who knew both cities, showed worse accuracy than German students who often knew only San Diego. In this and other experiments, people who knew more performed worse because the distal criterion—in this case population—was not known. The population of a city is correlated to its coverage in foreign newspapers, which in turn is correlated to city recognition in students: the higher the population, the more frequently foreign newspapers report about the city, and the higher is the recognition rate among students (Goldstein and Gigerenzer 2002). More recently, Hertwig et al. (2008) showed that population estimates for known cities are influenced by processing fluency, and the analysis we are going to present is consistent with these findings.

Let us assume German participants are asked whether San Antonio or San Diego has the larger population; if they recognize San Diego and cannot rely on another cue, they will answer that San Diego is the more populous city. Do these participants know that San Diego has a larger population? No, because knowledge means justified true belief: It could be that the opposite is true, and San Antonio is bigger.
Moreover, the participants might not know that recognition of a city is correlated with its population. Are these participants epistemically justified to believe that San Diego has a larger population? Yes, they are, because the use of the recognition heuristic yields a higher ratio of true beliefs than mere guessing, as suggested by Goldstein and Gigerenzer (2002). The authors acknowledged that the recognition heuristic is useful in some context, but not others, and Oppenheimer (2003) showed that participants discount recognition as cue to judge population if they know that recognition is caused by factors unrelated to size. An example is Chernobyl, which became famous for its nuclear accident.

How could people connect recognition to city size? Some participants may have experienced that city population goes with higher recognition. When reading about France, participants may have experienced that they recognize the larger cities (Paris, Marseille, Lyon), but not the smaller ones (Autun, Beaune, Nevers). Later, when reading about Poland, the same happens, and again when reading about Italy. Even if these participants do not remember the relationship between city population and city recognition, the two are associated through earlier experience. A person can implicitly establish the link between city population and recognition, which is later taken as an indicator of city size if recognition is not known to be due to other factors (e.g., Beaune for producing vine). This is essentially what Unkelbach (2007) found when studying participants who learned a fluency-truth link. Other participants in the study may not have had direct experience with recognition of cities and city size, but they may have experienced that recognition goes with greater importance or larger size in general. For example, people may have experienced that politicians they recognize hold more important positions than politicians they do not recognize, or that companies they recognize are larger than those they do not recognize.

Logically, taking recognition as a cue to city population is affirming the consequent because city size, mediated by media coverage, leads to recognition. However, in the absence of alternative cues to population, this can be seen as an inference to the best explanation (Harman 1965). Note that it is not clear from the studies of Goldstein and Gigerenzer (2002) whether participants knew about the link between city population and recognition; however, the appeal of the findings most plausibly stems from the fact that people normally do not explicitly know about such links. Therefore, use of the recognition heuristic is not part of deductive reasoning, but—if we are right about how the recognition heuristic emerged—about using a process to assess importance, or about learning connections between recognition of a stimulus and its size or importance.

After the illustration of justification in the case of the recognition heuristic, we are going to use the terminology introduced above in order to assess the epistemic status of judgments based on processing fluency that a belief is true.

3 Epistemic Justification of Fluency

The main question we try to answer is whether using processing fluency as a cue is a valid process to determine truth and therefore justifies a belief. Does processing fluency provide an above-chance probability of true beliefs? In order to answer the main question, we have to examine the probability that fluency indicates truth.
Although there is little empirical research into correlations between fluency and distal criterion in the environment, we can provide a quantitative analysis of the epistemic justification of fluency, given the proportion of fluent statements (with fluency as the proximal cue) that are judged as being true, and given the proportion of fluent statements that are actually true (distal criterion). We focus our analysis on repetition as a source for processing fluency. However, a similar analysis may be done for other sources of fluency, such as implicit learning or semantic coherence. In this section, we first present a Bayesian analysis, then outline the logic of epistemic justification of fluency and finally discuss why we assume that fluent statements are more likely to be true than false.

3.1 Bayesian Analysis

In the following, we consider only statements one has been exposed to; this means that these are repeated statements. We are not interested in the infinite number of possible statements, but in the huge, but finite number of statements a person has encountered before. This means that in the following Bayesian analysis, true will always indicate statements or propositions a person has encountered before (see also Footnote 5). In other words, we set $p(\text{exposure})=1$ for true statements and disregard statements—true or untrue—a person has not encountered yet.

We use Bayes’ Theorem to derive the probability that a fluent statement is true from the probabilities that a true statement becomes fluent by repeated exposure and the base rate of true statements, which here means the probability of the truth of a statement a person has encountered before:

$$p(\text{true}|\text{fluent}) = \frac{p(\text{fluent}|\text{true})p(\text{true})}{p(\text{fluent}|\text{true})p(\text{true}) + p(\text{fluent}|\neg\text{true})p(\neg\text{true})}$$

(1)

As the expression “$p(\text{fluent}|\text{true}) * p(\text{true})$” is in the denominator and the numerator, Formula (1) has the structure $Y = A/A + B$. If $A>B$, then $Y > .5$. As $Y$ is $p(\text{true}|\text{fluent})$, we have shown that one can derive $p(\text{true}|\text{fluent}) > .5$ from $p(\text{fluent}|\text{true})$ and $p(\text{true})$ if one assumes that $p(\text{fluent}|\text{true})p(\text{true}) > p(\text{fluent}|\neg\text{true})p(\neg\text{true})$. One assumption is that statements one has encountered before become fluent regardless of their truth status, i.e., $p(\text{fluent}|\text{true}) = p(\text{fluent}|\neg\text{true})^4$; every exposure to a statement has the same effect when people do not know the truth, and we can simplify Formula 1:

$$p(\text{true}|\text{fluent}) = \frac{p(\text{true})}{p(\text{true}) + p(\neg\text{true})}$$

(2)

Therefore, the outcome of the calculation depends solely on $p(\text{true})$ and $p(\neg\text{true})$: $p(\text{fluent}|\text{true}) * p(\text{true}) > p(\text{fluent}|\neg\text{true}) * p(\neg\text{true})$ if $p(\text{true}) > p(\neg\text{true})$, which amounts to the claim that $p(\text{true}|\text{fluent}) > .5$, as outlined above. This means that the probability that a statement is true given that it is fluent is more than .5 if a person

---

Footnote 4: This assumption of course may be wrong: For example, it seems plausible that repeated false statements are less fluent because they cohere less with pre-existing information stored in memory. However, there is no empirical evidence for this claim and it works in favor of our hypothesis, that is, if false statements were less fluent than true statements, fluency would be even more valid as a cue to truth than we propose here.
has encountered more true than false statements. As we assume that there are only true and false statements, then \( p(\text{true}) + p(\text{false}) = 1 \); therefore, a probability >.5 is above chance.\(^5\)

3.2 Fluency and Epistemic Justification

The standard we need in order to justify the use of fluency as a cue to truth is the proportion of accurate judgments, that is, the probability to correctly identify true and false statements, given that fluency is used as a cue. Mathematically, this probability is given by the probability of true statements being fluently processed, \( p(\text{true} | \text{fluent}) \), multiplied by the probability of judging fluently processed statement as being true, \( p(\text{judged "true" | fluent}) \), plus the probability of fluently processed false statements, \( p(\text{false} | \text{fluent}) \), multiplied by the probability of judging a fluently processed statement as false, \( p(\text{judged "false" | fluent}) \). As all these probabilities include the conditional probability that a statement is processed fluently, the probability of an accurate judgment is simplified to:

\[
p(\text{accurate}) = p(\text{true}) \cdot p(\text{judged "true" | fluent}) + p(\text{false}) \cdot p(\text{judged "false" | fluent})
\]

If we assume that there are only true and false statements, then \( p(\text{false}) = 1 - p(\text{true}) \); the same applies for judged truth and falsehood. Thus, the formula is (again, omitting the conditional probability of fluent processing):

\[
p(\text{accurate}) = p(\text{true}) \cdot p(\text{judged "true" | fluent}) + (1 - p(\text{true})) \cdot (1 - p(\text{judged "true" | fluent}))
\]

This formula incorporates some tacit assumptions: First, processing fluency for all repeated statements is the same; there is no gradient, for example, with exposure frequency (i.e., two exposures lead to more fluency than one exposure). Second, all relationships are linear. Third, we assume that the benefits of hits and correct rejections is the same as the costs of false alarms (saying “true” when it is false) and misses (saying “false” when it is true). Fourth, we assume that there are no other factors that influence the decision (factual knowledge, trustworthiness of the source, etc.); that is, the formula refers to truth judgments under uncertainty, in a similar way as the recognition heuristic refers to choices under uncertainty (see above). Fifth, as already mentioned, we are interested only in statements people have been exposed to before.

None of these assumptions are far-fetched or unrealistic, but represent the most conservative way of estimating the epistemic justification of fluency. As it is, Formula (4) shows that \( p(\text{accurate} | \text{fluent}) \), the probability to judge fluently processed statements accurately, only depends on \( p(\text{true} | \text{fluent}) \) and \( p(\text{judged "true" | fluent}) \). The latter is the probability that people use fluency in judgments of

\(^5\) It is again important to remember that the analysis includes only statements a person has encountered before. One reviewer made the following calculation when all possible statements were included: The base rate \( p(\text{true}) \) is \( 1/1000 \); \( p(\text{fluent | true}) = 1/3 \), and \( p(\text{fluent | false}) = 1/1000 \). Using Formula (1), this yields \( 1/3 \cdot 1/1000 + 1/1000 \cdot 999/1000 \cdot 1/3/3997/3000 = 1/4 \). This example demonstrates how much our analysis hinges on the base rate of true statements, and this base rate critically depends on which statements to include. As we look at fluency and truth from the viewpoint of an individual, our analysis is based on the question whether it is justified for an individual to believe in statements he or she has encountered before and are therefore processed fluently.
truth. As long as this probability is above 50%, people are using fluency as a cue. The former is the ecological correlation that is necessary for fluency to be epistemically justified. Let us now examine this relationship:

How can we now derive meaningful estimates for these probabilities? Empirically, Unkelbach and Stahl (2009, Exp. 1) found that this probability was about 60%. In this study, 71 students judged 120 statements as “true” or “false”. Half of these statements were factually true and half were factually false; however, because these were difficult statements, this factor had no influence on the results. More importantly, 60 statements were new, while participants had encountered the other half of the statements 30 min earlier in a presentation phase. Participants judged these old statements to be true with a probability of \( M = .587 \) (SD = .134), a value that is significantly different from 50%, \( t(70) = 5.37, p < .001 \); a non-significant result would have indicated that people are not using fluency as a cue at all. That is, on average, each participant judged about 60% of fluently processed statements as true. Thus, we will use this empirical value as a starting point.

One might argue that a rational decision maker should always use fluency (i.e., \( p(\text{judged } \text{“true”}|\text{ fluent}) = 1.0 \)) if the cue is valid. However, people do not adhere to rational choice in such situations, but rather follow a process of “probability matching” (Erev and Barron 2005). That is, if choice A (i.e., judging “true”) leads to an accurate outcome in 75% of all cases, they do not always choose A, but on average in only 75% of all cases. The adaptiveness of such behaviour is extensively discussed elsewhere (Erev and Barron 2005; but also, Shanks et al. 2002).

Therefore, for illustration, assume that people judge fluent statements “true” in 60% of the cases, and that 90% of the fluent statements are factually true (i.e., \( p(\text{judged } \text{“true”}|\text{ fluent}) \)): From 100 statements, 90 are true and 10 are false. 60% of both the 90 true statements and of the 10 false statements are judged to be true. This relationship is displayed in Table 1. The proportion of accurate answers is then:

\[
p(\text{accurate}|\text{fluent}) = 0.9 \times 0.6 + 0.1 \times 0.4 = 0.58
\]  

We assume that the proportion of “true” judgments does not depend on whether a statement a person has been exposed to is actually true or false: In our example, all these repeated statements should be judged as “true” because they are fluently processed due to prior exposure, regardless of their actual truth. In other words, we are not dealing with conditional probabilities. If we used a formula with conditional probabilities: \( p(\text{accurate}|\text{true}) = p(\text{judged “true”}|\text{true}) \times p(\text{true}) \), the result would be the same: \( p(\text{judged “true”}|\text{true}) = .54/.90 \) and \( p(\text{true}) = .90 \), so that this calculation yields .54; for the conditional probabilities of false statements, \( p(\text{accurate}|\text{false}) = p(\text{judged “false”}|\text{false}) \times p(\text{false}) \), with \( p(\text{judged “false”}|\text{false}) = .04/.10 \)

|                | True | False | Total |
|----------------|------|-------|-------|
| Judged “true”  | 54   | 6     | 60    |
| Judged “false” | 36   | 4     | 40    |
| Total          | 90   | 10    | 100   |

Table 1 Frequencies of judgments for true and false statements when 60% of repeated statements are judged true and 90% of the statements are true
and \( p \) (false) = .10; this yields .04, as given in Table 1. It is therefore simpler to calculate the unconditional probabilities. As noted earlier, the assumptions are simplified, but there is no empirical evidence that would justify the assumption that probabilities depend on whether a statement is true or false.

Obviously, the critical assumption in this estimation is that 90% of the fluently processed statements are true. Fig. 1 shows what happens to our indicator of justification, \( p \) (accurate | fluent), when different levels of the ecological correlation (i.e., \( p \) (true | fluent)) and cue usage (i.e., \( p \) (judged “true” | fluent)) are considered. This graph shows that as long as more than 50% of fluently processed statements are true, the usage of fluency is ecologically valid and thus epistemically justified, as this results in \( p \) (accurate | fluent) being at above chance level. The extreme examples illustrate this graph: when 100% of all judgments are true, 60% of all judgments will be correct. If none of the fluent statements were true, only 40% of the judgments would be correct.

3.3 Are Fluently Processed Statements More Likely to be True than False?

Why should \( p \) (true | fluent) be greater than 50%? Again, according to Bayes’ Theorem (see Formula 1), this probability is high when \( p \) (fluent | true) is high and when \( p \) (true) is high. Although there is no direct empirical evidence on the ecological probabilities of true and false statements people encounter, there are at least two reasons why we can speculate that fluently processed statements are more likely to be true rather than false, which means that \( p \) (fluent | true) is high.

First, true statements about facts where the distal criteria are known should have higher repetition probabilities and consequently, be processed with greater fluency, than false statements. As there is a large number of plausible false statements (e.g., “The present height of the Cheops-Pyramid is 123/124/xxx.m.”), derived from only one true statement (e.g., “The present height of the Cheops-Pyramid is 138.75 m.”), the true statement should be encountered more frequently. Unless a specific falsehood is established, true statements should prevail. As we discussed above, there is also no need for verbatim repetition, but structural similarity is sufficient (Reber 1967; Buchner 1994). This logical necessity should be amplified by Grice’s maxim of quality (see below; Grice 1975).

Second, if statements indeed reflect something about the true state of the world, a given set of statements should be coherent (e.g., “Venus is closer to the Sun than the
Earth.”—“The Earth is the third planet from the Sun.”—“Venus is the second planet from the Sun.”), as mentioned in our earlier discussion on coherence as a derivative standard (Goldman 1986). Thus, if a given set of statements is incoherent, then at least one statement cannot be true. Recently, Topolinski and colleagues observed that semantic coherence in word triads leads to greater processing fluency than incoherence, even when people are not aware of the coherence (Topolinski et al. 2009; Topolinski and Strack 2009a, b). Their experiment showed that people’s experience of fluent processing indeed is an indication of coherence while disfluency indicates factual incoherence. As disfluency is an indicator of incoherence, and incoherence implies lack of truth, we conclude that disfluent statements are less likely to be true than fluent statements.

In addition, $p$ (true) should also be high; as cooperative communicators assume that a speaker’s utterances are relevant (Sperber and Wilson 1986), and relevance includes that speakers say what they think is true (Grice’s 1975, maxim of quality), listeners are entitled to assume that the statements they hear are true. Preliminary evidence that $p$ (true) is high comes from research about the accuracy of encyclopedias, such as Wikipedia. Unfortunately, some studies focus on errors so that no measure of accuracy is given (e.g., Giles 2005). As only relatively few errors were found (see supporting material of Giles 2005, for a list of errors), one can assume that accuracy of these encyclopedia articles—and therefore $p$ (true) in this narrow set of statements—is high. One study that examined the accuracy of both Wikipedia and a peer-reviewed web site (Rajagopalan et al. 2010), the patient-oriented National Cancer Institute's Physician Data Query (PDQ) comprehensive cancer database, observed that inaccuracies were very rare on both sources (<2%).

Especially given the last point, which assumes that speakers usually communicate what they believe to be true, our illustrative analysis certainly underestimates $p$ (fluent | true), because we exclude all cases for which truth not only leads to fluent processing, but is also derived from explicit memory or, for example, syllogistic reasoning (Klauer et al. 2000). And there is clear experimental evidence that participants use both explicit remembering and non-referential sources to judge truth (Brown and Nix 1996). That is, if one remembers having encountered a statement, then one can assume that the statement is true, if it does not contradict the derivative standard of coherence among one’s own beliefs and knowledge. However, even the experience of explicit remembering may be guided by an experience of processing fluency (Kurilla and Westerman 2008; Mandler 1980); thus, fluent processing increases the experience of remembering a statement. As just outlined, a person who remembers a statement is entitled to assume that the statement is true. If fluency related to explicit remembering of a statement indicates that the statement is true, we can conclude that fluently processed statements that later are remembered are more probably true. The pure experience of processing fluency under exclusion of direct remembering or reasoning that we have depicted above therefore clearly underestimates the usefulness of the fluency cue.

In the past, the usage of fluency as a cue has often been treated as a judgmental bias or an error, because in most experimental setups, previous encounters and factual truth are artificially orthogonalized (e.g., Mitchell et al. 2005). Our analysis suggests (as illustrated in Fig. 1) that as long as more than half of the statements one has been exposed to are true, it is epistemically justified to use processing fluency of
a statement, derived from prior exposure (see Mandler 1980) and other possible sources, as a cue for the truth of the statement.

4 Fluency and Commonsense Beliefs

The assumption that we defend in this paper, that fluency is a valid cue to truth, may indicate how a universal mechanism of cognitive processing contributes to cultural diversity. The universal mechanism we postulate is that people interpret processing fluency in terms of the truth value of a statement. As fluency is positively correlated to the truth of a statement, a fluently processed statement is more likely to be interpreted as being true, even in the absence of evidence that supports its truth. For the sake of epistemological analysis, we have considered statements that refer to some distal criterion whose truth can be justified with relative ease, such as factual information from geography, history, politics, or science (i.e., it is easy to verify that cactuses cannot procreate by parthogenesis). We now are going to broaden the scope in that we consider the judged truth of commonsense beliefs. We first discuss some potential exceptions to our notion that it is justified to believe in statements that were encountered before and therefore can be processed fluently; we then turn to commonsense beliefs whose truth can only be determined with difficulty, if at all.

4.1 Some Exceptions

People learn early in life to distinguish between people who tell the truth and those who do not: Even pre-school children are able to identify and monitor sources of true and false statements and assign selective trust according to the credibility of the source (e.g., Koenig et al. 2004; Clément 2010). However, it has been shown that people not always are able to jointly represent the content of a statement and the credibility of the source and therefore commit the error to judge a familiar statement from an untrustworthy source as being true (e.g., Hovland and Weiss 1951). This was called the sleeper effect. More recently, it has been shown that people indeed use familiarity as a cue to truth even if they learned that the information was wrong (Brown and Nix 1996). Therefore, just identifying a trustworthy source is not sufficient to monitor the truth of information because people may forget the source of information before they forget the information itself, or they may remember both the information and its source, but dissociate the two. Thus, our assumption holds true only if most of the information is true, or if most of the sources of information are trustworthy.

Exceptions to the rule that most information is true or most sources of information are trustworthy readily come to mind in the context of persuasion, such as product marketing and political campaigns. In both cases, repetition is a crucial aspect. If false information is repeated (such as that some potato chips are healthy or that economic production increased), marketers and campaigners use—or abuse—the inclination of their addressees to assign truth to fluently processed information. If a source is untrustworthy, as in the study by Koenig et al. (2004), people may try to discount this source, but still risk that they fall prey to the sleeper effect. Further
research should investigate which strategies best help detect the fallacies lurking behind the processing of intentionally forged information.

Finally, in this study we only examined information a person has encountered before. We therefore cannot apply our analysis to statements that are true, but that people encounter for the first time. Of course, authoritarian regimes not only provide false information, but try to suppress the truth—exactly for the reason that people should not become familiar with it. Thus, our analysis suggests that a statement that can be fluently processed is more probably true than false, but it remains to be analyzed to what degree new statements that have been never encountered before indeed are false. Concluding that a belief in a fluently processed statement is justified does not entail the justification of disbelief in a new statement. Déchene and colleagues (2009a) as well as Unkelbach and Stahl (2009) found some evidence that new statements are not as easily believed; however, the effects were small in comparison to the advantages that familiar and fluent information enjoys. For a formal analysis of justification of disbelief in new statements, we would have to include all possible statements, not just those that a person has encountered before.

4.2 Common Sense Beliefs and Doxa

There are classes of opinion statements whose truth can be determined only with difficulty, if at all, because the external criteria are either unclear or absent. A good example of such statements is commonsense beliefs. They can be further divided into popular beliefs and unquestioned commonsense beliefs, called “doxa” by Bourdieu (1972/1977). An example may illustrate what the difference is: When assessing the accuracy of the statement “monarchy is a good form of government”, people living in monarchies may be more likely to endorse this form of government than people living in republics. This is a popular opinion, as people know about alternative forms of governments, but not a doxa as described by Bourdieu. An example for the latter is the discussion of the cities of the ladies by Christine de Pizan (1405/1982). Her discussion, for all its modernity in promoting women’s role in society, is completely entrenched in the monarchic structures of her time. At that time, monarchy was not an option among others; it was the only, unquestioned form of government in the world she knew. Would one have asked her whether she finds monarchy a good form of government, she probably would have been surprised at the very question: What would have been the alternative?

Despite the futility, or at least difficulty, of justifying the truth of such statements by reference to external criteria, the cognitive mechanisms by which people assess the accuracy of these types of statement presumably are the same as the ones with which they assess the truth of factual statements (see Rappaport 1999, 310, for a similar view). This may support the hypothesis that people generalize the positive correlation between fluency and truth learned on factual statements to commonsense beliefs. In contrast to factual statements, the judged truths of which people globally agree upon (e.g., “Paris is the capital of France”), commonsense beliefs are more culture-specific and often believed without further questioning. By hearing a commonsense belief or just living in a culture and getting exposed to its habits and customs, people process a specific set of beliefs more easily than alternative beliefs. We have seen that this experience of fluency increases judged truth of factual
statements, and it presumably increases the judged accuracy of statements that lack external criteria for epistemic justification. This does not mean that there is no way these statements could be justified. Popular opinions or doxastic beliefs may reinforce customs and behaviors that render commonsense beliefs accurate or at least adaptive within the respective culture, despite their lack of universal truth which can be ascribed to factual statements (see Boyer 1992; Rappaport 1999, for similar views on the role of truth in rituals).

A processing fluency account might then be able to explain how people justify their beliefs and doxa without clear external criteria: People, we submit, may assign truth to such “facts” simply because, as part of a given culture, they are extremely easy to process, while possible alternatives are new and difficult. Thus, processing fluency could be an internal cognitive mechanism for the perpetuation and justification of such otherwise difficult-to-verify statements and opinions.

5 Conclusion

We have reviewed empirical evidence that fluency, defined as the experienced ease of ongoing mental processes, influences judgments of truth: Fluently processed stimuli are judged to be true with higher probability than non-fluently processed stimuli. That is, processing fluency is used as a proximal cue for the distal criterion of truth. While this was regarded as a judgmental bias and error, we have shown for factual information that the use of fluency as a cue to truth is epistemically justified, if we make the assumption that most statements one is exposed to are true. Using mathematical models and empirical data, we showed that as long as the percentage of true factual statements that are fluently processed is above 50%, the use of fluency is justified. In correlational terms, we proposed a positive ecological correlation between fluency and truth and gave a number of reasons while we believe why this assumption may hold. For popular opinions and doxastic beliefs, we concluded, it is more difficult to account for epistemic justification on the basis of fluency.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

Arkes, H.R., C. Hackett, and L. Boehm. 1989. The generality of the relation between familiarity and judged validity. Journal of Behavioral Decision Making 2: 81–94.

Arkes, H.R., L.E. Boehm, and G. Xu. 1991. Determinants of judged validity. Journal of Experimental Social Psychology 27: 576–605.

Bacon, F.T. 1979. Credibility of repeated statements: Memory for trivia. Journal of Experimental Psychology: Human Learning and Memory 5: 241–252.

Begg, I.M., V. Armour, and T. Kerr. 1985. On believing what we remember. Canadian Journal of Behavioural Science 17: 199–214.

Begg, I.M., A. Anas, and S. Farinacci. 1992. Dissociation of processes in belief: Source recollection, statement familiarity, and the illusion of truth. Journal of Experimental Psychology: General 121: 446–458.
McGlone, M.S., and J. Tofighbakhsh. 2000. Birds of a feather flock conjointly (?): Rhyme as reason in aphorisms. *Psychological Science* 11: 424–428.

Mitchell, J., C.S. Dodson, and D.L. Schacter. 2005. fMRI evidence for the role of recollection in suppressing misattribution errors: The illusory truth effect. *Journal of Cognitive Neuroscience* 17: 800–810.

Oppenheimer, D.M. 2003. Not so fast! (and not so frugal!): rethinking the recognition heuristic. *Cognition* 90: B1–B9.

Pascal, B. (1678/1941). *Pensées* (W. F. Trotter, Trans.). New York: Modern Library.

Quine, W.V. 1969. *Epistemology naturalized*. Ontological relativity and other essays. New York: Columbia University Press.

Rajagopalan, M.S., V. Khanna, M. Stott, Y. Leiter, T.N. Showalter, A. Dicker, and Y.R. Lawrence (2010). Accuracy of cancer information on the Internet: A comparison of a Wiki with a professionally maintained database. *Journal of Clinical Oncology*, 28 (7 s), abstr 6058.

Rappaport, R.A. 1999. *Ritual and religion in the making of humanity*. Cambridge: Cambridge University Press.

Rawson, K.A., and J. Dunlosky. 2002. Are performance predictions for text based on ease on processing? *Journal of Experimental Psychology: Learning, Memory, and Cognition* 28: 69–80.

Reber, A.S. 1967. Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior* 6: 855–863.

Reber, R., and N. Schwarz. 1999. Effects of perceptual fluency on judgments of truth. *Consciousness and Cognition* 8: 338–342.

Reber, R., N. Schwarz, and P. Winkielman. 2004. Processing fluency and aesthetic pleasure: Is beauty in the perceiver’s processing experience? *Personality and Social Psychology Review* 8: 364–382.

Roediger, H.L., and K.B. McDermott. 1995. Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory and Cognition* 21: 803–814.

Rubin, D.C. 1995. *Memory in oral traditions: The cognitive psychology of epic, ballads, and counting-out rhymes*. Oxford: Oxford University Press.

Shanks, D., R. Tunney, and J. McCarthy. 2002. A re-examination of probability matching and rational choice. *Journal of Behavioral Decision Making* 15: 233–250.

Skurnik, I., N. Norbert Schwarz, and P. Winkielman. 2000. Drawing inferences from feelings: The role of naive beliefs. In *The message within: The role of subjective experience in social cognition and behavior*, ed. H. Bless and J.P. Forgas, 162–175. Philadelphia: Psychology.

Sperber, D., and D. Wilson. 1986. *Relevance. Communication and cognition*. Oxford: Blackwell.

Topolinski, S., K.U. Likowski, P. Weyers, and F. Strack. 2009. The face of fluency: Semantic coherence automatically elicits a specific pattern of facial muscle reactions. *Cognition and Emotion* 23: 260–271.

Topolinski, S., and F. Strack. 2009a. The analysis of intuition: Processing fluency and affect in judgements of semantic coherence. *Cognition and Emotion* 23: 1465–1503.

Topolinski, S., and F. Strack. 2009b. The architecture of intuition: Fluency and affect determine intuitive judgments of semantic and visual coherence, and of grammaticality in artificial grammar learning. *Journal of Experimental Psychology: General* 138: 39–63.

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: 219–230.

Unkelbach, C., and C. Stahl. 2009. A multinomial modeling approach to dissociate different components of the truth effect. *Consciousness & Cognition* 18: 22–38.

Whittlesea, B.W.A. 1993. Illusions of familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19: 1235–1253.

Whittlesea, B.W.A., and L.D. Williams. 1998. Why do strangers feel familiar, but friends don’t? A discrepancy-attribute account of feelings of familiarity. *Acta Psychologica* 98: 141–165.

Whittlesea, B.W.A., and L.D. Williams. 2001. The discrepancy-attribute hypothesis: II. Expectation, uncertainty, surprise, and feelings of familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 27: 14–33.

Winkielman, P., and N. Schwarz. 2001. How pleasant was your childhood? Beliefs about memory shape inferences from experienced difficulty of recall. *Psychological Science* 12: 176–179.