Please cite as: Scholl, S., Greifeneder, R., & Bless, H. (2014). When fluency signals truth: Prior successful reliance on fluency moderates the impact of fluency on truth judgments. *Journal of Behavioral Decision Making, 27*, 268–280. doi: 10.1002/bdm.1805

When Fluency Signals Truth: Prior Successful Reliance on Fluency Moderates the Impact of Fluency on Truth Judgments

Sabine G. Scholl[1], Rainer Greifeneder[2], and Herbert Bless[1]

[1]University of Mannheim, Germany [2]University of Basel, Switzerland

Word count: 11.737 words

Author Note

Sabine G. Scholl (née Czenna), Department of Psychology, University of Mannheim; Rainer Greifeneder, Department of Psychology, University of Basel; Herbert Bless, Department of Psychology, University of Mannheim

This research was supported by a grant from the Deutsche Forschungsgemeinschaft within the SFB 504 of the University of Mannheim, Germany, and a post-doctoral grant from the Landesstiftung Baden-Württemberg awarded to the second author. We are grateful to Annette Bauer, Stephan Bedenk, Marcia Duriska, Mariela Jaffé, Anne Landhäußer, Sarah Lohmüller, Mareike Wickop, Marleen Rusche, and Katharina Zimmer for their help in data collection.

Correspondence concerning this article should be addressed to Dr. Sabine G. Scholl (née Czenna), School of Social Sciences, University of Mannheim, 68131 Mannheim, Germany. E-mail: sabine.scholl@uni-mannheim.de; Phone: +49 (0)621 181-1975; Fax: +49 (0)621 181-2038

Abstract

Repeated statements are more frequently judged to be true. One position relates this so-called "truth effect" to meta-cognitive experiences of fluency, suggesting that repeated statements are more frequently judged to be true because they are processed more fluently. While most prior research focused on *why* repetition influences truth judgments, considerably less is known about *when* fluency is used as information. The present research addresses this question and investigates whether reliance on fluency is moderated by learning experiences. Specifically, we focus on changes in the reliance on fluency over the course of time. A series of experiments reveals that fluency is more likely to be used in truth judgments when previous reliance on fluency has resulted in valid judgments, compared to when previous reliance on fluency was misleading. These findings suggest that reliance on fluency in judgments is a finely-tuned process that takes prior experiences with fluency-based judgments into account.

Keywords: fluency, truth effect, feedback learning, subjective experiences

When Fluency Signals Truth: Prior Successful Reliance on Fluency Moderates the Impact of Fluency on Truth Judgments

Prior research has demonstrated that repeated statements are rated more frequently as true than new statements (e.g., Arkes, Hackett, & Boehm, 1989; Hasher, Goldstein, & Toppino, 1977; Hawkins & Hoch, 1992). This effect was denoted truth effect (see Hasher, et al., 1977, for a first description; Schwartz, 1982) and has been reliably replicated across research domains. One prominent explanation holds that repeated statements are rated more frequently as true because, compared to new statements, they can be processed more fluently (Reber & Schwarz, 1999). Specifically, it is assumed that the fluency experienced while encoding a statement is interpreted and used as a metacognitive cue with respect to the statement's truth. Corroborating this fluency-truth hypothesis, it has been demonstrated that increasing processing fluency independent of repetition, also produces truth effects. For instance, it has been shown that statements printed in high as compared to low color contrast are processed more fluently, and that this processing fluency translates into increased ratings of truth (e.g., Reber & Schwarz, 1999; Unkelbach, 2007). Moreover, statements are rated as more probably true when prior semantic activation (priming) facilitates processing (Kelley & Lindsay, 1993). Because fluency is manipulated independently of repetition, both of these findings suggest that repetition is not a necessary condition for truth effects. Rather, it appears that repetition-based truth effects are mediated via processing fluency (cf., Reber & Schwarz, 1999; see Unkelbach, Bayer, Alves, Koch, & Stahl, 2011; for a discussion of the independent effects of fluency and positivity).

Yet why do individuals associate fluency with truth, and not with falsehood? Schwarz (2004) proposed that individuals hold naïve theories about what their mental processing means, and use these naïve theories to interpret experienced fluency. Naïve theories thus likely constitute the link between fluency experiences and subsequent judgments, and

different naïve theories may give rise to different inferences and conclusions. For example, Winkielman and Schwarz (2001) asked participants to recall events from their childhood. Subsequently, participants' naïve theories on whether positive or negative events are more difficult to retrieve were manipulated. The ease with which events from childhood could be retrieved influenced participants' evaluations of their childhood—however, this impact depended on whether participants believed that positive or negative childhood events come to mind more easily, thus exemplifying the notion that the same metacognitive experience can give rise to very different conclusions. Intriguingly, naïve theories of meaning are themselves a function of prior learning. Unkelbach (2007) proposed that the fluency-truth link is learned throughout ontogenesis via feedback and reflects contingencies in the environment, such that true statements have a higher probability of being repeated and can thus be processed more fluently. In support of this argument, Unkelbach (2007) reported evidence showing that individuals can learn to associate fluency with falsehood (and not truth) by pairing fluent processing with falsehood. Note that explanations of naïve theories and feedback learning are not in competition but complement each other.

Most research on the truth effect has concentrated on establishing the effect, that is, on *whether*, *how*, and in *which direction* fluency influences judgments of truth. Much less is known about *when* individuals use fluency as a signal of truth (and when they do not) and most importantly, how reliance on fluency changes over the course of time due to learning. The present research addresses these questions and suggests that individuals rely on fluency experiences especially when this strategy has proven successful in the past. We thus propose that individuals not only learn "what" to infer from fluency (Unkelbach, 2007), but also "whether" fluency provides a judgmental basis for successful inferences in the first place. By way of this, the present contribution further advances our understanding of why there may be "wisdom in feelings" (Schwarz, 2002).

Moderators of fluency effects in judgments of truth

Research on moderators of the truth effect can be categorized into two groups: truth effects that result from repetition, and truth effects that are independent of repetition. While both categories of findings focus on the same effect, namely the influence of more or less fluent processing on judgments of truth, their moderators differ strongly and have received unequal attention in previous research. With respect to repetition, a recent meta-analysis identified several moderators such as gist versus verbatim repetition, delay between sessions 1 and 2, presentation of mixed lists containing old and new statements versus only one type of statement, response format, modality of experiment, modality of data collection, and level of processing (Dechêne, Stahl, Hansen, & Wänke, 2010). Additionally, other studies identified declining memory (Brown & Nix, 1996), declining context memory with age (Skurnik, Yoon, Park, & Schwarz, 2005), source credibility (Begg, Anas, & Farinacci, 1992), source variability (Roggeveen & Johar, 2002), and the direction of the repetition-truth relation (Unkelbach & Stahl, 2009) as moderators. Interestingly, these moderators pertain exclusively to repetition-based truth effects as they investigate different ways of inducing or experiencing repetition. To the best of our knowledge, there are only two studies that directly investigate which variables moderate fluency effects on judgments of truth that are independent of repetition. Specifically, Koch and Forgas (2012) showed that negative mood eliminated the truth effect compared to positive or neutral mood. Research by Hansen, Dechêne, and Wänke (2008) reveals that perceptual fluency affects judgments of truth when the discrepancy between actual and expected fluency is high and not low (for an overview of moderators of other fluency effects, see Alter & Oppenheimer, 2009; Greifeneder, Bless, & Pham, 2011; see also Whittlesea & Williams, 2000).

Against the background of the rather limited evidence on moderators of repetition-independent truth effects, the present research investigates the flexibility of individuals'

reliance on fluency experiences over the course of time when forming judgments of truth. We propose that reliance on fluency depends on the situational context and, more specifically, that judgments of truth reflect fluency experiences particularly when prior reliance on fluency has been a successful strategy. Although no direct evidence on this question is available, at least four lines of research pertaining to the influence of subjective experiences are relevant.

First, it has been shown that when individuals think about the reasons for specific subsequent judgments, they rely less on fluency-based information sources such as familiarity than when individuals were not instructed to think about reasons (Halberstadt & Catty, 2008). Presumably this is because when asked to give reasons, individuals have a preference for rationally justifiable information sources, and are unable or not willing to use their feelings.

Second, research suggests that individuals cease to rely on fluency experiences when the experiences' diagnosticity is called into question. On the one hand, diagnosticity can be questioned by explicitly providing a reason for increased or decreased fluency. Specifically, Schwarz and colleagues (1991) reported that ease-of-retrieval effects on assertiveness judgments—the fluency with which information is retrieved from memory—were eliminated when experienced retrieval fluency was attributed to a source unrelated to the judgmental target. Similarly, Wänke, Schwarz, and Bless (1995) showed that frequency judgments were determined by the perceived diagnosticity of retrieval fluency (see also Briñol, Petty, & Tormala, 2006; for a review, see Schwarz, 1998; Winkielman, Schwarz, & Belli, 1998, Greifeneder, et al., 2011). On the other hand, diagnosticity can be questioned more implicitly. For instance, individuals were shown to spontaneously discount and overcompensate for the perceived influence of fluency on judgments of frequency when alternative explanations for increased or decreased fluency were salient (Oppenheimer, 2004). Note that, in these cases, fluency experiences were rather *explicitly* dislinked from the judgmental target, that is,

fluency experiences were no longer attributed to the target stimulus. However, what if fluency is attributed to the target, but has proven unhelpful in prior learning (trials)?

Third, indirect evidence for a flexible use of fluency that depends on the perceived validity of prior feeling-based judgments has been reported in the realm of affective feelings. Specifically, research by Avnet, Pham, and Stephen (2012) demonstrated that individuals rely more on their current mood state when they believed they had been successful with past reliance on their mood state rather than when they believe that they had been unsuccessful.

Finally, Unkelbach (2007) has demonstrated that the interpretation of fluency in truth judgments is a function of its ecological validity (see also Unkelbach, 2006). When fluency was previously coupled with truth, participants exhibited the repetition-independent truth effect; however, when fluency was previously coupled with falsehood, fluent as compared to non-fluent presentation resulted in less (and not more) perceived truth. These findings suggest that individuals may adaptively learn the direction of the fluency-truth-correlation from prior experiences. Note that Unkelbach (2006; 2007) focused on the direction of the information provided by fluency. This manuscript focuses on learned validity and most importantly, on changes in reliance on fluency over the course of time.

The present research

Going beyond prior research, we propose that individuals are sensitive to whether prior reliance on fluency has resulted in correct or incorrect judgments. We argue that individuals' reliance on fluency in truth judgments is a finely tuned metacognitive process that takes prior learning into account, thus guarding against undue inferences. Accordingly, individuals should tend to trust fluency experiences when prior fluency-based judgments were correct, whereas they should be cautious about relying on fluency experiences when prior judgments were incorrect. The focus of this manuscript is on changes in the reliance on fluency over the course of time.

To test these considerations, we conducted five experiments in which participants evaluated the truth of a large number of trivia statements. Building on prior research (e.g., Reber & Schwarz, 1999; Reber, Winkielman, & Schwarz, 1998; Unkelbach, 2007), fluency was manipulated by means of color contrast. To manipulate prior success or failure when relying on fluency, participants received feedback indicating that reliance on fluency either was or was not a successful strategy. The proportion of judgments consistent with a fluency-based judgmental strategy served as the main dependent variable. If participants do not rely on fluency, this measure should be at chance-level. In contrast, if participants rely on fluency as a signal of truth, this measure should be above chance-level. Across the five experiments, we tested (a) the hypothesis that individuals rely on fluency in truth judgments when fluency previously proved to be a successful cue (Experiments 1 to 5), (b) whether this effect generalizes to situations in which no feedback is provided (Experiments 2, 3, and 4), (c) whether this effect holds up when the similarity between learning and test trials is reduced (Experiment 3), or when the source of fluency is different between learning and test trials (Experiment 4), and (d) whether this effect disappears when other cues that do not differ in fluency are provided (testing pattern matching as an alternative hypothesis in Experiment 5).

Because our experiments are built upon the methodology introduced by Unkelbach (2007), we start by reviewing Unkelbach's approach more closely. Unkelbach (2007) provided strong evidence that the direction of the influence of fluency on truth judgments is a function of prior learning. While he investigated whether the direction of fluency can be learned, his hypotheses were not geared towards changes in reliance on fluency over the course of time as investigated in the present manuscript. In light of this difference in focus, two aspects of the employed methodology deserve discussion. Both aspects concern methodological choices with respect to the *learning phase*.

First, Unkelbach (2007) used only easy statements (e.g., "The formula for water is H_2O."). Because participants may be expected to have *known* the answers to at least some of these statements, participants may have used fluency and prior knowledge as cues in judgment. A judgment that looks like reliance on fluency does therefore not necessarily reflect reliance on fluency. Note that this is not problematic when fluency and truth status are orthogonal (as in the control condition in the learning phase of Unkelbach, 2007, Exp. 2) and when signal detection measures are used for analyses, because prior knowledge about truth status and the response bias' magnitude can be estimated independently.[1] It is, however, less diagnostic when fluency and truth status are perfectly correlated (as in the classic and reversed condition in the learning phase of Unkelbach, 2007, Exp. 2), because even signal detection analyses then only allow for overall estimates, but not for specific estimates for the high versus low fluent statements (based on which the truth effect is calculated). Note that this is unproblematic for the research question pursued by Unkelbach, but is critical when investigating the change in reliance on fluency in the learning phase and from the learning to the test phase, as pursued in the present project. For that reason, we changed Unkelbach's methodology such that in the learning phase only difficult statements were used. Note that using only pretested difficult statements reduces the likelihood that participants know the truth but cannot forestall it, thus creating noise in the data. Obtaining an effect, however, would attest to the strength of the effect of feedback learning. Additionally, using statements where prior knowledge is likely helpful (Unkelbach, 2007) and not (present experiments) allows for more general conclusions about the learned reliance on fluency experiences.

Second, by coupling fluency and truth status, two different sources of fluency were intermixed in Unkelbach's (2007, Exp. 2) learning phase. Specifically, prior research has shown that truth/false status in the real world may be an independent source of fluency (Gilbert, 1991; Unkelbach, 2007; Unkelbach & Stahl, 2009). Fluency due to color contrast

may thus be enhanced or decreased by fluency due to truth status (Unkelbach, 2007; Unkelbach & Stahl, 2009), resulting in a situation where the relative difference between high versus low color contrast statements is larger when high fluency is paired with truth (one condition), and smaller when high fluency is paired with falsehood (other condition). This is not problematic when investigating whether the meaning of fluency can be changed and when the focus is placed on the test phase only (as done by Unkelbach). It is, however, less diagnostic when the reliance on fluency *during* the learning phase and *across* learning and test phase is the main focus of interest as in the present experiments. For that reason, we used factually true statements only.

These considerations are important, because as a consequence of these features, Unkelbach (2007) confined conclusions to the primary goal of demonstrating the learned *direction* of the information provided by fluency—even though a control group was available that resembles, on the surface level, the "fluency-not-valid groups" in the present experiments. However, for the reasons detailed above, our fluency-not-valid groups allow for testing whether the reliance on fluency changes over the course of time.

In sum, building on Unkelbach's (2007) finding that the direction of the information provided by fluency (true or false) can be learned, we investigate whether individuals may learn over the course of time to take fluency's past validity into account when forming truth judgments. To this end, small but critical changes in methodology were introduced that allow for more precise conclusions about whether and to what extent fluency was used in truth judgments in the different phases of the experiment.

Experiment 1

In a feedback-learning task, participants were asked to judge the truth of a series of factually true trivia statements differing in color contrast: Some statements were presented in high color contrast (high readability and associated high fluency), and other statements in low

color contrast (low readability and associated low fluency). Participants received tailored feedback about the correctness of their judgments so that some participants would learn that reliance on fluency is a successful strategy (readability-valid), while other participants would learn that fluency is not related to truth or falsehood (readability-not-valid). We expected participants to react to the feedback and to rely on fluency in the readability-valid condition, that is, to judge statements more frequently as true when presented in high as compared to low fluency. In contrast, participants in the readability-not-valid condition were expected to refrain from relying on fluency because the feedback implied that reliance on fluency is not a successful strategy. Importantly, this should be particularly apparent in later trials of the experiment, because learning from repeated feedback presumably requires numerous learning experiences.

Method

Eighty-eight students of a German university (43 females; M_{age} = 23.6 years, SD = 3.4) were randomly assigned to a 2 (readability high vs. low) x 2 (readability-valid vs. readability-not-valid) x 2 (phase 1 vs. 2) mixed factorial design with readability-manipulation and phase as within factors.

Participants read 56 factually true statements drawn from encyclopedia resources that were of similar length and were pre-tested as highly ambiguous with respect to truth (we chose statements for which 40% to 60% of participants indicated that they were true and for which certainty ratings were below 3 on a 5-point scale, ranging from 1 = absolutely uncertain to 5 = absolutely certain; sample statements read “The latest eruption of Láscar volcano was on 18th April 2006;” “Zirkalloy is characterized by high neutron permeability”). The order of statements was determined randomly but was identical for all participants. As the hypothesized effect was expected to emerge in later trials—that is, after participants had the chance to react to the provided feedback—the 56 statements were divided in two sets of 28

statements each, which will subsequently be referred to as Phase 1 and Phase 2. Readability was manipulated within participants by presenting statements in one of eight color contrasts (foreground-to-background), resulting in either high levels of contrast and associated perceptual fluency (e.g., dark green/light olive-green, dark purple/light purple), or low levels of contrast and associated perceptual fluency (e.g., olive-green/light olive-green, dark purple/purple). Each statement was presented individually and remained on the screen until ”true/false” judgments were made. Participants provided their answers by clicking on ”true” or ”false” buttons, while response latencies were recorded.[2] Subsequently, on a new screen (without seeing the statement any more), participants received feedback about the correctness of this answer. This feedback was tailored so that participants in the *readability-valid* condition would learn a positive relationship between fluency and ostensible truth (r = 1.0). In contrast, participants in the *readability-not-valid condition* would not learn such a relationship, because in this condition half of the high-readable and half of the low-readable statements were said to be true, whereas the remaining statements were said to be false, thus rendering readability undiagnostic about truth (r = .0). Both groups were not informed about how many high and low readable statements were said to be true or false. After receiving feedback, participants started the next trial. Finally, participants were debriefed and informed that all statements were true, were probed for suspiciousness, thanked, and remunerated with 2 EUR.

Results and Discussion

For Phase 1 and 2, we calculated how often participants' judgments were consistent with a fluency-based judgmental strategy (high readability= true; low readability= false). While participants in the readability-valid condition indicated that 61.3% of the high readable statements and 51% of the low readable statements were true, participants in the readable-not-valid condition indicated that 52.7% of the high readable statements and 58.5% of the low

readable statements were true. Specifically, the number of high-readable statements rated as true plus the number of low-readable statements rated as false was related to the total of 28 statements in each phase, yielding a percentage measure varying between 0 (all judgments contrary to reliance on fluency) and 100 (all judgments congruent with reliance on fluency), with 50 percent reflecting the chance-level midpoint. We opted for a percentage measure because it allows investigating changes in reliance on fluency over the course of time.

The percentage measures for Phase 1 and Phase 2 were subjected to two analyses: first, to compare the use of fluency between conditions and over the course of time a 2 (readability-valid vs. readability-not-valid) x 2 (phase 1 vs. 2) ANOVA with phase as within-factor was computed; second, to assess the use of fluency in comparison to chance-level, planned comparisons were calculated. Figure 1 shows the predicted interaction between condition and phase, $F(1, 86) = 6.15$, $MSE = 106.29$, $p < .02$, $\eta_p^2 = .07$. Participants in the readability-valid condition were more likely to rely on their fluency experiences in Phase 2 than in Phase 1 ($M = 51.95\%$, $SD = 13.81$, $M = 58.36\%$, $SD = 19.38$, respectively), $F(1, 43) = 7.22$, $MSE = 125.35$, $p < .02$, $d = .14$; whereas no increase was observed in the readability-not-valid condition ($M = 47.73\%$, $SD = 9.84$, $M = 46.43\%$, $SD = 7.47$), $F < 1$.

Planned comparisons with chance-level further underscore these findings. Participants in the readability-valid condition relied on fluency at chance-level in Phase 1 ($M = 51.95\%$, $SD = 13.81$), $t < 1$, and above chance-level in Phase 2 ($M = 58.36\%$, $SD = 19.38$), $t(43) = 2.86$, $p < .01$, $d = .43$. In contrast, participants in the readability-not-valid condition relied on fluency at chance-level in Phase 1 ($M = 47.73\%$, $SD = 9.84$), $t(43) = -1{,}53$, $p > .14$, and below chance-level in Phase 2 ($M = 46.43\%$, $SD = 7.47$), $t(43) = -3.17$, $p < .01$, $d = .48$. Overall, participants in the readability-valid condition relied more often on fluency experiences than participants in the readability-not-valid condition, $F(1, 86) = 11.34$,

MSE = 253.07, $p < .01$, $\eta_p^2 = .12$. The main effect for phase was not significant, $F < 2.71$, $p > .11$.[3]

The results suggest that participants were affected by the provided feedback. Participants relied on encoding fluency when forming truth judgments if prior feedback implied that such reliance is associated with correct judgments, but not when the feedback implied no relationship between fluency and truth. At first glance, a percentage level of marginally over 55 may not appear impressive. However, the observed effects are of similar or higher magnitude than those generally reported in investigations on the (repetition-independent) effect of fluency on truth judgments (e.g., Reber & Schwarz, 1999).

Experiment 2

Experiment 2 was designed to replicate and extend the findings of Experiment 1. Specifically, after the first two phases (with feedback), a third phase was added in which participants were no longer provided with feedback about the correctness of their judgments. If participants in the readability-valid condition learn in Phases 1 and 2 that fluency is a valid signal of truth, they should continue to rely on fluency even if no further feedback is provided in Phase 3.

Method

Seventy students from a German university (31 females; M_{age} = 22.1 years, *SD* = 2.7) were randomly assigned to a 2 (readability high vs. low) x 2 (readability-valid vs. readability-not-valid) x 3 (phase 1 vs. 2 vs. 3) mixed factorial design with readability manipulation and phase as within-factors. In comparison to Experiment 1, a new set of factually true statements was pre-tested and color contrasts were changed so as to eliminate the possibility that effects were due to specific color contrast combinations. Participants were presented with 30 statements in each of the three phases. During Phases 1 and 2, procedures were identical to

Experiment 1. Prior to Phase 3, participants were informed that the same task would continue, however without further feedback.

Results and Discussion

As in Experiment 1, the number of statements for which participants' judgments were consistent with a fluency-based judgmental strategy was related to the total of 30 statements in each phase.[4] These three percentage measures were analyzed with a planned Helmert contrast (2 -1 -1) (0 1 -1) in a 2 (readability-valid vs. readability-not-valid) x 3 (phase 1 vs. 2 vs. 3) mixed factorial design with phase as within-factor. This analysis results in two main effects of the contrast for phase and in two interactions of the contrast with condition. The interaction between the first effect of the contrast (2 -1 -1) and condition allows for testing the hypothesis that reliance on fluency increases from Phase 1 to Phases 2 and 3 for participants in the readability-valid condition, whereas no such increase is expected in the readability-not-valid condition. The second effect of the contrast (0 1 -1) and its interaction with condition allow for testing an orthogonal set of contrasts and investigating whether reliance on fluency changes between Phase 2 and 3 in general and whether this change differs between conditions. For phase, the first main effect of the contrast (2 -1 -1) revealed that, overall, participants relied less on fluency in Phase 1 than in Phases 2 and 3 (M_1 = 52.10%, SD_1 = 10.48 vs. M_2 = 54.57%, SD_2 = 16.75, M_3 = 57.62%, SD_3 = 15.96), $F(1, 68) = 5.03$, $MSE = 222.86$, $p < .03$ $\eta_p^2 = .07$. Additionally, as indicated by the second main effect of the contrast (0 1 -1), participants by tendency relied less on fluency in Phase 2 compared to Phase 3, $F(1, 68) = 3.33$, $MSE = 195.40.35$, $p = .08$, $\eta_p^2 = .05$. Simple effects analyses revealed that participants in the readability-valid condition relied more on fluency in Phases 2 and 3 than in Phase 1, $F(1, 34) = 6.19$, $MSE = 251.31$, $p < .02$, $\eta_p^2 = .15$, whereas this was not the case for participants in the readability-not-valid condition, $F < 1$. In both conditions, the reliance on

fluency did not differ between Phases 2 and 3, Fs < 2.36, ps > .14, MSEs > 251.31. Figure 2 shows the observed pattern of results.

Comparing the reliance on fluency with chance-level further underscores the differences between readability-valid and readability-not-valid conditions. Planned contrasts showed that in the readability-not-valid condition, reliance on fluency remained at chance-level in all three phases (M_1= 50.67%, SD_1= 9.35, M_2= 50.86%, SD_2 = 10.86, M_3 = 53.14%, SD_3 = 11.05), ts < 1.68, *ns.* In contrast, in the readability-valid condition, reliance on fluency was marginally above chance-level in Phase 1 (M = 53.52%, SD = 11.46), $t(34) = 1.82$, $p = .08$, $d = .31$, and reliably above chance-level in Phase 2 (M = 58.29%, SD = 20.57), $t(34) = 2.38$, $p < .03$, $d = .40$; importantly, fluency remained above chance-level in Phase 3 (M = 62.10%, SD = 18.81), $t(34) = 3.81$, $p <.01$, $d = .64$. Additionally, overall, participants in the readability-valid condition relied more often on fluency than participants in the readability-not-valid condition (M = 57.97%, SD = 14.53, M = 51.56%, SD = 6.06), $F(1, 68) = 5.81$, $MSE = 123.85$, $p < .02$, $\eta_p^2 = .08$. In sum, the observed pattern supports the hypothesis that individuals particularly use fluency in truth judgments when prior reliance was successful. It should be noted, however, that the predicted first interaction of the contrast between phase and readability-valid/-not-valid-condition failed to reach conventional levels of significance, $F < 2.23$, $MSE = 222.86$, $p = .14$.

The results of Experiment 2 replicate and extend the findings observed in Experiment 1. Again, when participants could learn that fluency is a valid signal of truth, reliance on fluency increased and was significantly above chance-level. In contrast, if the feedback was tailored such that fluency was not associated with truth, reliance on fluency remained at chance-level. Extending Experiment 1, this effect generalized to subsequent situations in which no further feedback about the correctness of participants' responses was provided. These findings attest that individuals adaptively rely on fluency when forming truth

judgments depending on the success that was associated with reliance on fluency in preceding trials.

Experiment 3

In combination, Experiments 1 and 2 support the hypothesis that when forming truth judgments, individuals rely on the fluency of encoding as a function of whether or not such reliance has previously proven successful. Alternatively, one may speculate that participants did not rely on fluency, but based their judgments on learned associations between certain color-combinations and the truth status of statements. Note however, that in each experiment, we not only used one but several *different* color contrasts to manipulate high versus low fluency, rendering it unlikely that participants learned to associate specific combinations (e.g., dark olive-green/light olive-green) with truth or falsehood. Nevertheless, it appeared desirable to replicate and extend Experiment 2, such that in the third phase, fluency is manipulated differently from the first two phases (for a similar approach, see Unkelbach, 2007). To alienate Phase 3 from Phases 1 and 2—and thus to reduce the likelihood that individuals apply a specific color contrast = truth rule—different sets of color contrast were used for Phases 1 and 2 versus Phase 3. Additionally, statements were slightly tilted in Phase 3 to reduce presentation similarity to Phases 1 and 2 on a surface level. We expected that the generalization effect which emerged in Phase 3 of Experiment 2 would hold up even in the new alienated Phase 3. This expected pattern would question alternative explanations that rest on learned associations other than the hypothesized fluency-truth link.

Method

Forty-eight students from a German university (24 females; M_{age} = 23.3 years, *SD* = 3.1) were randomly assigned to a 2 (readability high vs. low) x 2 (readability-valid vs. readability-not-valid) x 3 (phase 1 vs. 2 vs. 3) mixed factorial design with readability and phase as within-factors. Manipulations and dependent variables were identical to Experiment

2, apart from the following changes: In Phase 3, besides providing no feedback, fore- and background-colors were different from Phases 1 and 2. Additionally, statements were unsystematically slightly tilted to render Phase 3 more different from Phases 1 and 2 on the surface level.

Results and Discussion

The percentage of participants' judgments consistent with a fluency-based judgmental strategy was calculated for each of the three phases. These percentage measures were analyzed using a planned Helmert contrast (2 -1 -1) (0 1 -1) in a 2 (readability-valid vs. readability-not-valid) x 3 (phase 1 vs. 2 vs. 3) mixed factorial design with phase as within-factor. As predicted, the first interaction of the contrast (2 -1 -1) with condition was significant, $F(1, 46) = 7.86$, $MSE = 213.74$, $p < .01$, $\eta_p^2 = .15$, and reflected in further analyses: Figure 3 shows that from Phase 1 to Phases 2 and 3, the reliance on fluency increased significantly for participants in the readability-valid condition, $F(1, 23) = 4.78$, $MSE = 278.02$, $p < .04$, $\eta_p^2 = .17$, and decreased, as a non-significant tendency, for participants in the readability-not-valid condition, $F(1, 23) = 3.10$, $MSE = 149.45$, $p = .10$, $\eta_p^2 = .12$.

Planned comparisons with chance-level further underscore these findings: While participants in the readability-valid condition relied on fluency at chance-level in Phase 1 ($M_1 = 53.27\%$, $SD_1 = 15.33$), $t < 1$, and above chance-level in Phases 2 and 3 ($M_2 = 59.97\%$, $SD_2 = 18.30$, $M_3 = 61.46\%$, $SD_3 = 18.36$), all ts > 2.67, ps < .02, ds > .54; participants in the readability-not-valid condition relied on fluency at chance-level in all three phases ($M_1 = 54.02\%$, $SD_1 = 11.60$, $M_2 = 48.51\%$, $SD_2 = 8.49$, $M_3 = 50.74\%$, $SD_3 = 12.14$), $|t$s$| < 1.70$, ps > .11. Neither the second interaction of the contrast (0 1 -1) between phase and condition, nor the two main effects of the contrast for phase were significant, Fs < 1.

The results of Experiment 3 further support the hypothesis that individuals adaptively use fluency in truth judgments. Replicating Experiments 1 and 2, participants were more likely to rely on the fluency with which a statement could be encoded when reliance on fluency has proven successful in preceding situations.

Experiment 4

Experiments 1 to 3 support the hypothesis that individuals rely on fluency as a function of whether reliance on fluency previously proved to be successful or not. This was the case even when Phase 3 was alienated from Phases 1 and 2, suggesting that the observed pattern of results does not hinge on simple learned associations between specific color contrasts and truth. While Experiment 3 refutes some alternative explanations, one may still wish for more evidence attesting to the hypothesized fluency-truth link. Particularly, as participants in Experiment 3 could still learn that contrast *in general* is indicative for truth. In order to rule out this alternative interpretation and to show that learning effects still occur when using different kinds of fluency manipulations, we introduced another change to Phase 3. Specifically, by manipulating fluency in a way different from that used in Phases 1 and 2, we created a situation in which Phases 1 and 2 as compared to Phase 3 are only conceptually similar (all manipulate fluency), but vary notably in terms of operationalization. Should participants' judgmental behavior still carry over from Phases 1 and 2 to Phase 3, this would strongly support the suggested fluency-truth link, because alternative feature-based explanations (e.g., associations between color contrasts and truth) are not viable.

Method

Fifty students from a German university (41 females; M_{age} = 23.0 years, SD = 5.6) were randomly assigned to a 2 (readability high vs. low) x 2 (readability-valid vs. readability-not-valid) x 3 (phase 1 vs. 2 vs. 3) mixed factorial design with readability and phase as within-factors. Manipulations and dependent variables were identical to Experiments 2 and 3,

apart from the following important changes: In Phase 3, readability was manipulated by reducing the space between letters and by rotating statements by 180° (in two-dimensional space). Specifically, high readable statements were presented with normal letter spacing (100%) and non-rotated, whereas low readable statements were presented with reduced letter spacing (70%) and rotation. To create variation, four different font-types of similarly high legibility were used (Times New Roman, Arial, Calibri, and Cambria).

Results and Discussion

For each of the three phases, the percentage of participants' judgments consistent with a fluency-based judgmental strategy was calculated. These percentage measures were analyzed using a planned Helmert contrast (2 -1 -1) (0 1 -1) in a 2 (readability-valid vs. readability-not-valid) x 3 (phase 1 vs. 2 vs. 3) mixed factorial design with phase as within-factor. Again, the first interaction of the contrast (2 -1 -1) with condition was significant, $F(1, 48) = 5.19$, $MSE = 180.09$, $p < .03$, $\eta_p^2 = .10$. As can be seen in Figure 4, reliance on fluency increased from Phase 1 to Phases 2 and 3 for participants in the readability-valid condition, $F(1, 24) = 3.96$, $MSE = 272.57$, $p = .06$, $\eta_p^2 = .14$, but did not change for participants in the readability-not-valid condition, $F(1, 24) = 1.22$, $MSE = 87.61$, $p > .28$. Neither the second interaction of the contrast (0 1 -1) between phase and condition, nor the two main effects of the contrast for phase were significant, $Fs < 1.41$, $MSEs > 314.35$, $ps > .25$. Overall, participants in the readability-valid-condition relied more on fluency than participants in the readability-not-valid-condition, $F(1, 48) = 12.74$, $MSE = 156.27$, $p < .001$, $\eta_p^2 = .21$.

Planned comparisons with chance-level further underscore these findings: While participants in the readability-valid condition relied on fluency above chance-level in all three phases ($M_1 = 56.29\%$, $SD_1 = 14.30$, $M_2 = 65.00\%$, $SD_2 = 24.24$, $M_3 = 60.71\%$, $SD_3 = 21.63$), all $ts > 2.20$, $ps < .04$, $ds > .44$; participants in the readability-not-valid condition relied on

fluency at chance-level in all three phases (M_1 = 49.43%, SD_1 = 7.34, M_2 = 47.29%, SD_2 = 10.53, M_3 = 47.43%, SD_3 = 8.22), $|t|s < 1.57$, ps > .14.

The results of Experiment 4 further support the hypothesis that individuals adaptively rely on fluency when forming truth judgments. Replicating Experiments 1 to 3, participants were more likely to rely on fluency when reliance on fluency has previously proven successful. Critically, this pattern emerged even though fluency was elicited differentially in Phases 1 and 2 versus Phase 3, thus reducing the link between the learning phases and the test phase to the conceptual level. That the effect of reliance on fluency carried over even in such conditions strongly attests to the suggested fluency-truth link and refutes potential alternative explanations.

Experiment 5

Experiments 1 to 4 support the hypothesis that individuals rely on fluency experiences depending on whether reliance on fluency previously proved to be successful or not. This was even the case when fluency elicited in Phase 3 compared to Phases 1 and 2 differed on a surface-level (Experiment 3) or was only similar on the conceptual level (Experiment 4). Still, one could argue that what participants learnt is not a specific fluency association as argued here, but a general pattern linking one cue (e.g., fluency) to another cue (e.g., truth), and that participants simply matched this pattern from Phases 1 and 2 to Phase 3. This alternative account of pattern matching allows for the speculation that any new cue introduced in Phase 3 would result in similar findings than the fluency cue of Experiments 2 to 4, because participants would simply match the pattern (one cue signals truth) to the new situation in Phase 3.[5] To test this alternative account, we changed Phase 3 so that statements were presented with an orange or blue frame. Note that only the frame but not the statements themselves were of different colors. This created a situation in which a new (fluency-unrelated) cue was present in Phase 3 to which the presumably learned general pattern (one

cue signals truth) could be matched. Should participants' judgments differ as a function of readability-condition in Phase 3 compared to Phase 1, this would support a general pattern matching account and question a specific fluency hypothesis. In contrast, if participants' judgments differentially reflect their learning experiences in Phases 1 and 2, but do not differ between Phases 1 and 3, then this would question a general pattern matching account and provide additional support for the suggested specific fluency-truth-link.

In addition to testing the pattern matching account, Experiment 5 was meant to allow for additional insights as to how participants experienced the learning phases. To this end we asked participants several questions about how they perceived the feedback and how they formed their judgments.

Method

Forty-three students from a German university (24 females; M_{age} = 22.0 years, *SD* = 2.7) were randomly assigned to a 2 (readability high vs. low) x 2 (readability-valid vs. readability-not-valid) x 3 (phase 1 vs. 2 vs. 3) mixed factorial design with readability and phase as within-factors. Manipulations and dependent variables were identical to the previous experiments, apart from the following changes: statements in Phase 3 did not differ in fluency (all were presented in black Arial letters on a white background) but differed in whether they were surrounded by an orange or blue frame. Additionally, participants were asked to indicate in an open answer format the thoughts they had during the learning phases, and their agreement with the following six statements pertaining to the learning phases (1= I do not agree; 7 = I fully agree): "I have received more positive feedback in some domains than in others." "The feedback was clear enough." "The feedback was sufficient enough." "I have known the truth status of most statements." "I had to guess the truth status of most statements." "For many statements, I have questioned the validity of the feedback."[6]

Results and Discussion

For each of the three phases, the percentage of participants' judgments consistent with a fluency/cue-based judgmental strategy was calculated. These percentage measures were analyzed using a planned simple contrast with Phase 1 as reference point (-1 1 0) (-1 0 1) in a 2 (readability-valid vs. readability-not-valid) x 3 (phase 1 vs. 2 vs. 3) mixed factorial design with phase as within-factor. As predicted, the main effect of the first simple contrast (-1 1 0), $F(1, 41) = 3.92$, $MSE = 126.87$, $p = .06$, $\eta_p^2 = .09$, was qualified by an interaction between the first simple contrast and condition, $F(1, 41) = 4.74$, $MSE = 126.87$, $p < .04$, $\eta_p^2 = .10$. As can be seen in Figure 5, reliance on fluency deceased in the readability-not-valid condition from Phase 1 to Phase 2, $F(1, 21) = 7.33$, $MSE = 153.06$, $p < .02$, and remained the same in the readability-valid-condition, $F(1, 20) = 0.02$, $MSE = 99.37$, $p = .88$. Importantly, suggesting that participants' judgmental behavior in Phase 3 did not differ from their behavior in Phase 1, the main effect of the second simple contrast (-1 0 1), $F(1, 41) = 1.14$, $MSE = 186.52$, $p = .29$, and the interaction between the second simple contrast and condition were non-significant, $F(1, 41) = 0.17$, $MSE = 186.52$, $p = .68$. Additionally, the main effect for condition was non-significant, $F(1, 41) = 0.03$, $MSE = 28.91$, $p = .86$.

Planned comparisons with chance-level further underscore these findings: Participants in the readability-not-valid condition relied on fluency at chance-level in Phase 1 ($M_1 = 52.44\%$, $SD_1 = 7.09$; $t(21) = 1.61$, $p = .12$) and below chance-level in Phase 2 ($M_2 = 45.29\%$, $SD_2 = 9.76$; $t(21) = -2.26$, $p < .04$, $d = 0.48$) , while participants in the readability-valid condition relied on fluency at chance-level in both phases ($M_1 = 49.66\%$, $SD_1 = 8.74$, $M_2 = 50.00\%$, $SD_2 = 8.53$), $|ts| < 0.18$, $ps > .86$. In Phase 3, participants' reliance on fluency in the readability-valid and the readability-not-valid-conditions was at chance-level ($M_3 = 48.30\%$, $SD_3 = 9.76$, $M_3 = 49.35\%$, $SD_3 = 9.90$), $|ts| < 0.80$, $ps > .43$.

Participants' open comments and their agreement with the six statements were analysed to obtain more insight in how participants experienced the learning phases. Suggesting a successful selection of difficult statements in Phases 1 and 2, 21 participants (48%) wrote that they did not know the truth status of most of the statements and that the statements were very difficult. Additionally, on a scale ranging from 1 to 7, participants more strongly agreed with having *guessed* (M = 6.30, SD = 0.94) than with having *known* the truth status of most of the statements (M = 1.81, SD = 1.14, $F(1, 41) = 239.86$, MSE = 1.80, $p < .0001$, $\eta_p^2 = .85$. Neither the main effect for condition nor the interaction between question and condition were significant, Fs < 0.27, ps > .60. Suggesting that participants did not consciously identify or apply a "high readability/fluency=true, low readability/fluency=false-strategy," none of the participants in the readability-valid-condition mentioned testing or applying such a strategy, and only one participant in the readability-not-valid condition mentioned having wondered whether color contrast was indicative of truth. Analyses further revealed that readability-valid and readability-not-valid conditions did not differ in whether they judged the feedback as sufficient (M = 3.43, SD = 1.89; M = 3.64, SD = 1.81) and clear (M = 6.00, SD = 1.55; M = 5.64, SD = 1.33), and whether they indicated having received more positive feedback in some knowledge areas than in others (M = 4.10, SD = 1.76; M = 4.18, SD = 1.53),$|t|$s < 1 . Most importantly, suggesting that participants were not suspicious with respect to the feedback provided, none of the participants mentioned having received false feedback and three participants explicitly stated that they were interested in the feedback. Furthermore, readability-valid and readability-not-valid conditions did not differ significantly in their responses to an item assessing whether they questioned the validity of the feedback obtained (M = 3.57 SD = 1.96; M = 3.95, SD = 2.17), $|t| < 1$. To further analyse the potential role of feedback questioning, we divided the sample at the median (Md = 4) and reanalyzed the data in a 2 (readability high vs. low) x 2 (readability-

valid vs. readability-not-valid) x 3 (phase 1 vs. 2 vs. 3) x 2 (questioning feedback below 4 vs. equal and above the median) mixed factorial design with readability and phase as within-factors. Pattern and significance levels remain the same as above. The main effect of the first simple contrast (-1 1 0), $F(1, 39) = 3.07$, $MSE = 127.08$, $p < .09$, $\eta_p^2 = .07$, was qualified by an interaction between the first simple contrast and condition, $F(1, 39) = 4.87$, $MSE = 127.08$, $p < .04$, $\eta_p^2 = .11$. Reliance on fluency decreased in the readability-not-valid condition from Phase 1 to Phase 2, $F(1, 20) = 6.38$, $MSE = 158.80$, $p < .02$, and remained the same in the readability-valid-condition, $F(1, 19) = 0.14$, $MSE = 93.69$, $p = .71$. Again, suggesting that participants' judgmental behavior in Phase 3 did not differ from their behavior in Phase 1, the main effect of the second simple contrast (-1 0 1), $F(1, 39) = 1.49$, $MSE = 189.23$, $p = .23$, and the interaction between the second simple contrast and condition were non-significant, $F(1, 39) = 0.27$, $MSE = 189.23$, $p = .61$. Most importantly, the main effect of questioning the feedback and its' interactions and all other effects were non-significant, Fs < 1.68 and ps > .20. In sum, this analysis showed that participants who questioned the feedback *more* do not behave significantly differently from participants who questioned the feedback *less*. Though at least partially informative, a median-split analysis cannot fully correct for the effects of questioning feedback. To further address this critical issue, we focused data analysis on those participants who hardly questioned the feedback (= below 4; $N = 18$), in a 2 (readability-valid vs. readability-not-valid) x 3 (phase 1 vs. 2 vs. 3) mixed factorial design with phase as within-factor. While being particularly conservative, this analysis procedure comes with the downside of a small sample size; we counteracted this downside by computing the error-term on the basis of the full design. Paralleling the above reported results, reliance on fluency tended to decrease in the readability-not-valid condition from Phase 1 to Phase 2, $F(1, 41) = 2.19$, $MSE = 126.87$, $p = .15$, and tended to remain the same in the readability-valid-condition, $F(1, 41) = 1.12$, $MSE = 126.87$, $p = .30$. Most importantly, the

analysis revealed the predicted interaction between the first simple contrast and validity-condition, $F(1, 41) = 3.22$, $MSE = 126.87$, $p = .08$. Again, suggesting that participants' judgmental behavior in Phase 3 did not differ from their behavior in Phase 1, the main effect of the second simple contrast (-1 0 1), $F(1, 41) = 2.19$, $MSE = 186.52$, $p = .15$, and the interaction between the second simple contrast and condition were non-significant, $F(1, 41) = 0.55$, $MSE = 186.52$, $p = .46$. Additionally, the main effect for the first simple contrast, $F(1, 41) = 0.09$, $MSE = 126.87$, $p = .77$, and the main effect for condition, $F(1, 41) = 0.02$, $MSE = 28.91$, $p = .88$, were non-significant. In sum, the median split analysis and the particularly conservative analysis with the reduced sample indicate that questioning the feedback does not significantly alter the general pattern of results. Very cautiously we therefore suggest that feedback questioning does not qualify as a strong alternative explanation.

The results of Experiment 5 provide additional support for the hypothesis that individuals adaptively rely on fluency when forming truth judgments. Importantly, judgmental behavior in Phases 1 and 3 did not differ, thus reducing the likelihood that the results are due to a potential alternative cue-based explanation. Participants' open comments and their agreement with the six statements suggest that participants neither engaged in pattern–matching nor deliberately identified a "high readability/fluency=true, low readability/fluency=false-strategy." Additionally, the adaptive use of fluency still holds when controlling for the influence of questioning the validity of the feedback obtained and when analysing only those participants who hardly questioned the feedback's validity. Contrary to Experiments 1 to 4, reliance on fluency did not increase for the readability-valid condition in Phase 2 compared to Phase 1. As a result, one might ask whether learning occurred in Experiment 5. Note, however, that Phase 2 was similar to the previous four experiments, which all reflected significant increases from Phase 1 to 2; it therefore appears possible that

the non-significant increase in Experiment 5 is due to chance. What is more important is the observed decrease in reliance on fluency in the readability-not-valid condition, suggesting that participants did learn from experiences in Experiment 5, too.

General Discussion

Previous research that investigated the impact of encoding fluency on truth judgments has focused primarily on establishing and documenting the effect, namely, on *whether, how,* and *in which direction* fluency influences truth judgments (e.g., Reber & Schwarz, 1999; Unkelbach, 2007). While there is considerable evidence about variables that moderate repetition-based truth effects (for an overview, see Dechêne, et al., 2010), much less is known about what moderates truth effects that are due to differences in encoding fluency but that are not due to repetition (see Hansen, et al., 2008, for an exception pertaining to the discrepancy between expected and experienced fluency). The results reported in the present contribution directly address this issue in various aspects.

Replicating previous research (e.g., Reber & Schwarz, 1999; Unkelbach, 2007), five experiments support the assumption that encoding fluency influences truth judgments. Specifically, statements were more likely to be considered true when the statements could be encoded fluently rather than non-fluently. Going beyond demonstrating that fluency-based truth effects are a reliable phenomenon, the present work shows that individuals' reliance on fluency is moderated by the experienced validity of fluency as a cue over the course of time. Specifically, participants' judgments were more likely to reflect an impact of encoding fluency when participants had previously experienced that relying on fluency resulted in correct judgments. Moreover, fluency as a cue received more weight when individuals had more learning trials to experience fluency as a valid cue for truth judgments. In contrast, when the feedback was tailored such that fluency was not associated with truth, reliance on fluency remained at chance-level or decreased over the course of time. Experiments 2 to 4 further

showed that reliance on fluency carried over from trials with feedback to trials in which no feedback was provided. The results of Experiments 3 to 5 suggest that the observed findings are not easily attributed to explanations other than fluency, such as an association between truth status and color contrast or a general pattern matching account. For instance, fluency continued to influence truth judgments even when manipulations of fluency were related only on a conceptual level (Experiment 4). The results of Experiment 5 further suggest that the adaptive use of fluency is independent of whether participants questioned the validity of the feedback obtained.

Three aspects of the present experiments deserve additional discussion. First, on the basis of the data, we cannot know for sure whether participants in the fluency-not-valid condition reduced reliance on fluency because participants had learned that fluency was a useless cue or because of a lack of learning experiences that fluency may be a valid cue. However, research by Unkelbach (2007) suggests that fluency is an ecological valid cue, and research by Oppenheimer (2004) holds that individuals rely on fluency experiences even in one-shot decisions. It therefore appears fair to assume that participants in the current set of studies started out by relying on fluency, increased their reliance when learning that fluency is a valid cue, and decreased their reliance when learning that fluency is not a valid cue.

Second, against the background of the here presented evidence, one may wonder whether relying on fluency is a more conscious or more unconscious strategy. The experiments reported here allowed for some hint as to this intriguing question by analyzing self-report data. Yet, despite the insights offered, it will likely prove fruitful in future research to develop methods that go beyond self-reports. This is because prompting participants for their strategy in think-aloud protocols or in post-hoc assessments (such as in Experiment 5) may not allow for valid conclusions. For instance, participants in Experiment 5 did not mention having identified or applied a “high readability/fluency=true, low

readability/fluency=false-strategy," though their judgmental behaviour strongly suggest that they did in Phase 2. Only one participant in the fluency-not-valid condition mentioned having tested for an association between color contrast and truth (but could not identify any). The more abstract concept fluency was not mentioned at all. This suggests that participants were not able to consciously name the true causes of their judgmental behaviour, and is reminiscent of Nisbett and Wilson's (1977) more general account on the problems associated with introspection. But even if participants were able to correctly name their judgmental strategy, would they dare saying that they relied on a feeling (instead of "hard facts"), given that reliance on feelings is generally not considered a rational or even adaptive strategy (Elster, 1999)? Problems with introspection and social desirability effects may thus reduce the validity of self-reports, generating a need for methods that allow for dissociating strategies by other means.

Third, the obtained findings demonstrate that individuals learn over the course of time that it may be more or less successful to base truth judgments on fluency. Specifically, Experiments 2 to 4 consistently show that participants' reliance on fluency reflects prior experiences with fluency as a cue to truth. Such a flexible use of fluency experiences is in line with general theorizing on the role of subjective feelings, which holds that feelings are a generally valuable and functional source of information (see Greifeneder, et al., 2011; Schwarz, 2002; Greifeneder, Bless, & Scholl, 2013). Unkelbach (2007), for instance, argued that individuals learn in everyday life that fluency and truth are positively correlated; hence that fluency is an ecologically valid source of information. Relatedly, Hertwig, Herzog, Schooler, and Reimer (2008) proposed that fluency is an ecologically valid cue in inferences of city sizes.

Interestingly, besides demonstrating that individuals rely on fluency when it is a valid source of information, the obtained results also suggest that individuals may adapt their

reliance on fluency to different situations. In the present set of studies, individuals' reliance on experienced fluency reflected prior learning experiences even in highly different contexts in which fluency experiences were similar on a conceptual level only. Bearing in mind the dangers of comparing data across different experiments, it appears instructive to note that the difference in the reliance on fluency between fluency-valid and fluency-not-valid conditions decreases with decreasing similarity between the feedback phases (1 and 2) and Phase 3. This places emphasis on individuals' sensitivity not only to prior learning experiences with fluency's validity but also to differences in the context in which truth judgments have to be made. The latter also converges with research on subjective experiences that has documented that individuals may refrain from relying on subjective experiences as a judgmental cue. For example, individuals in Unkelbach (2007) did not rely on fluency when they have learned that fluency is an invalid cue for truth judgments. Moreover, individuals refrained from relying on the ease-of-retrieval heuristic when they believed that the fluent versus non-fluent retrieval of information from memory was not related to the information that was retrieved (Greifeneder & Bless, 2007; Schwarz, et al., 1991). Note however, that none of these studies investigated reliance on feelings over the course of time as in the present manuscript.

The present evidence suggests that reliance on fluency in particular, and perhaps on feelings in general, is a generally sensible judgmental mechanism. To some, this conclusion may be surprising, given that the mechanisms underlying reliance on feelings in judgment (attribution to target, inferences based on naïve theories, see Schwarz, 2004) are often frugal and crude. The present findings offer further evidence to this puzzle by attesting to the flexibility with which individuals rely on fluency in truth judgments. In that respect, we speculate that most likely, the adaptive use of fluency is not limited to truth judgments. Indeed, because fluency has been shown to influence a large variety of judgments, including judgments of fame (Jacoby, Kelley, Brown, & Jasechko, 1989), knowing (Koriat, 1993),

policy support (Garcia-Marques & Mackie, 2001), and frequency (Aarts & Dijksterhuis, 1999; Tversky & Kahneman, 1973), and evaluations of inferences (Ozubko & Fugelsang, 2011), it appears fair to assume that the here observed adaptive use of fluency likely extends to other types of judgments, too.

In concluding, it is interesting to note that learning associations between fluency and truth may not only regulate the direction of what to infer from a specific fluency experience (as shown in previous research, Unkelbach, 2007), but also when to infer meaning from a specific fluency experience in the first place. This is important because the judgmental consequences of "having some meaning" are likely different from "no meaning." When fluency means "something," fluency may be a reasonable source of information in judgment; when fluency has no meaning with respect to a specific target, it should not be used at all. As such, the present findings extend earlier evidence and attest to their critical role of previous experiences with the judgmental validity of fluency as a cue in fluency research.

References

Aarts, H., & Dijksterhuis, A. (1999). How often did I do it? Experienced ease of retrieval and frequency estimates of past behavior. *Acta Psychologica, 103*(1-2), 77-89.

Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the tribes of fluency to form a metacognitive nation. *Personality and Social Psychology Review, 13*(3), 219-235.

Arkes, H. R., Hackett, C., & Boehm, L. (1989). The generality of the relation between familiarity and judged validity. *Journal of Behavioral Decision Making, 2*(2), 81-94.

Avnet, T., Pham, M. T., & Stephen, A. T. (2012), Consumers' trust in feelings as information. *Journal of Consumer Research*, 39, 720-735.

Begg, I. M., Anas, A., & Farinacci, S. (1992). Dissociation of processes in belief: Source recollection, statement familiarity, and the illusion of truth. *Journal of Experimental Psychology: General, 121*(4), 446-458.

Briñol, P., Petty, R. E., & Tormala, Z. L. (2006). The malleable meaning of subjective ease. *Psychological Science, 17*(3), 200-206.

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

Dechêne, A., Stahl, C., Hansen, J., & Wänke, M. (2010). The truth about the truth: A meta-analytic review of the truth effect. *Personality and Social Psychology Review, 14*(2), 238-257.

Elster, J. (1999). *Alchemies of the mind: Rationality and the emotions*. Cambridge: Cambridge University Press.

Fazio, R. H. (1990). A practical guide to the use of response latency in social psychological research. In C. Hendrick & M. S. Clark (Eds.), *Research methods in personality and social psychology*. (pp. 74-97). Thousand Oaks, CA US: Sage Publications, Inc.

Gilbert, D. T. (1991). How mental systems believe. *American Psychologist, 46*(2), 107-119.

Greifeneder, R., & Bless, H. (2007). Relying on accessible content versus accessibility experiences: The case of processing capacity. *Social Cognition, 25*(6), 853-881.

Greifeneder, R., Bless, H., & Pham, M. T. (2011). When do people rely on affective and cognitive feelings in judgment? A Review. *Personality and Social Psychology Review, 15*(2), 107-141.

Greifeneder, R., Bless, H., & Scholl, S. G. (2013). About swift defaults and sophisticated safety nets: A process perspective on fluency's validity in judgment. In C. Unkelbach & R. Greifeneder (Eds.), *The experience of thinking* (pp. 220-233). Hove, UK: Psychology Press.

Halberstadt, J., & Catty, S. (2008). Analytic thought disrupts familiarity-based decision making. *Social Cognition, 26*(6), 755-765.

Hansen, J., Dechêne, A., & Wänke, M. (2008). Discrepant fluency increases subjective truth. *Journal of Experimental Social Psychology, 44*(3), 687-691.

Hasher, L., Goldstein, D., & Toppino, T. (1977). Frequency and the conference of referential validity. *Journal of Verbal Learning and Verbal Behavior, 16*(1), 107-112.

Hawkins, S. A., & Hoch, S. J. (1992). Low-involvement learning: Memory without evaluation. *Journal of Consumer Research, 19*(2), 212-225.

Hertwig, R., Herzog, S. M., Schooler, L. J., & Reimer, T. (2008). Fluency heuristic: A model of how the mind exploits a by-product of information retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(5), 1191-1206.

Jacoby, L. L., Kelley, C. M., Brown, J., & Jasechko, J. (1989). Becoming famous overnight: Limits of the ability to avoid unconscious influences of the past. *Journal of Personality and Social Psychology, 56*, 326-338.

Kelley, C. M., & Lindsay, D. S. (1993). Remembering mistaken for knowing: Ease of retrieval as a basis for confidence in answers to general knowledge questions. *Journal of Memory and Language, 32*(1), 1-24.

Koch, A. S., & Forgas, J. P. (2012). Feeling good and feeling truth: The interactive effects of mood and processing fluency on truth judgments. *Journal of Experimental Social Psychology, 48*(2), 481-485.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*(3), 231-259.

Oppenheimer, D. M. (2004). Spontaneous discounting of availability in frequency judgment tasks. *Psychological Science, 15*(2), 100-105.

Ozubko, J. D., & Fugelsang, J. (2011). Remembering makes evidence compelling: Retrieval from memory can give rise to the illusion of truth. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(1), 270-276.

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

Reber, R., Winkielman, P., & Schwarz, N. (1998). Effects of perceptual fluency on affective judgments. *Psychological Science, 9*(1), 45-48.

Reber, R., Wurtz, P., & Zimmermann, T. D. (2004). Exploring 'fringe' consciousness: The subjective experience of perceptual fluency and its objective bases. *Consciousness and Cognition, 13*(1), 47-60.

Roggeveen, A. L., & Johar, G. V. (2002). Perceived source variability versus familiarity: Testing competing explanations for the truth effect. *Journal of Consumer Psychology, 12*(2), 81-91.

Schwartz, M. (1982). Repetition and rated truth value of statements. *The American Journal of Psychology, 95*(3), 393-407.

Schwarz, N. (1998). Accessible content and accessibility experiences: The interplay of declarative and experiential information in judgment. *Personality and Social Psychology Review, 2*(2), 87-99.

Schwarz, N. (2002). Situated cognition and the wisdom in feelings: Cognitive tuning. In L. F. Barrett & P. Salovey (Eds.), *The wisdom in feeling: Psychological processes in emotional intelligence* (pp. 144-166). New York, NY, US: Guilford Press.

Schwarz, N. (2004). Metacognitive experiences in consumer judgment and decision making. *Journal of Consumer Psychology, 14*(4), 332-348.

Schwarz, N., Bless, H., Strack, F., Klumpp, G., Rittenauer-Schatka, H., & Simons, A. (1991). Ease of retrieval as information: Another look at the availability heuristic. *Journal of Personality and Social Psychology, 61*(2), 195-202.

Skurnik, I., Yoon, C., Park, D. C., & Schwarz, N. (2005). How warnings about false claims become recommendations. *Journal of Consumer Research, 31*(4), 713-724.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments & Computers, 31*(1), 137-149.

Unkelbach, C. (2006). The learned interpretation of cognitive fluency. *Psychological Science, 17*(4), 339-345.

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., Bayer, M., Alves, H., Koch, A., & Stahl, C. (2011). Fluency and positivity as possible causes of the truth effect. *Consciousness and Cognition: An International Journal, 20*(3), 594-602.

Unkelbach, C., & Stahl, C. (2009). A multinomial modeling approach to dissociate different components of the truth effect. *Consciousness and Cognition: An International Journal, 18*(1), 22-38.

Wänke, M., Schwarz, N., & Bless, H. (1995). The availability heuristic revisited: Experienced ease of retrieval in mundane frequency estimates. *Acta Psychologica, 89*(1), 83-90.

Whittlesea, B. W. A., & Williams, L. D. (2000). The source of feelings of familiarity: The discrepancy-attribution hypothesis. *Journal of Experimental Psychology: Learning, Memory and Cognition, 26*(3), 547-565.

Winkielman, P., & Schwarz, N. (2001). How pleasant was your childhood? Beliefs about memory shape inferences from experienced difficulty of recall. *Psychological Science, 12*(2), 176-179.

Winkielman, P., Schwarz, N., & Belli, R. F. (1998). The role of ease of retrieval and attribution in memory judgements: Judging your memory as worse despite recalling more events. *Psychological Science, 9*(2), 124.

Wurtz, P., Reber, R., & Zimmermann, T. D. (2008). The feeling of fluent perception: A single experience from multiple asynchronous sources. *Consciousness and Cognition, 17*(1), 171-184.

Figure Captions

Figure 1. Means (with standard errors) of participants' reliance on fluency in Experiment 1, separately for the readability-valid condition (n = 44) versus the readability-not-valid condition (n = 44), and the two experimental phases.

Figure 2. Means (with standard errors) of participants' reliance on fluency in Experiment 2, separately for the readability-valid condition (n = 35) versus the readability-not-valid condition (n = 35), and the three experimental phases (Phases 1 and 2 with feedback; Phase 3 without feedback).

Figure 3. Means (with standard errors) of participants' reliance on fluency in Experiment 3, separately for the readability-valid condition (n = 24) versus the readability-not-valid condition (n = 24), and the three experimental phases (Phases 1 and 2 with feedback; Phase 3 without feedback).

Figure 4. Means (with standard errors) of participants' reliance on fluency in Experiment 4, separately for the readability-valid condition (n = 25) versus the readability-not-valid condition (n = 25), and the three experimental phases (Phases 1 and 2 with feedback; Phase 3 without feedback).

Figure 5. Means (with standard errors) of participants' reliance on fluency/general cue in Experiment 5, separately for the readability-valid condition (n = 21) versus the readability-not-valid condition (n = 22), and the three experimental phases (Phases 1 and 2 with feedback; Phase 3 without feedback and general but not fluency cue).

Figure 1

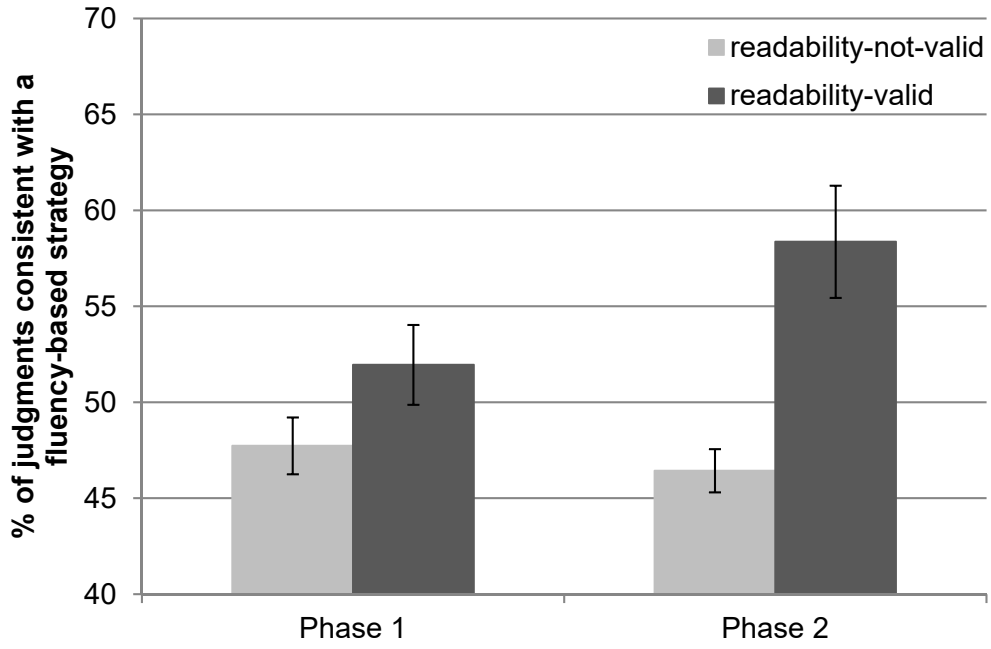

70
65
60
55
50
45
40
% of judgments consistent with a fluency-based strategy
readability-not-valid
readability-valid
Phase 1
Phase 2

Figure 2

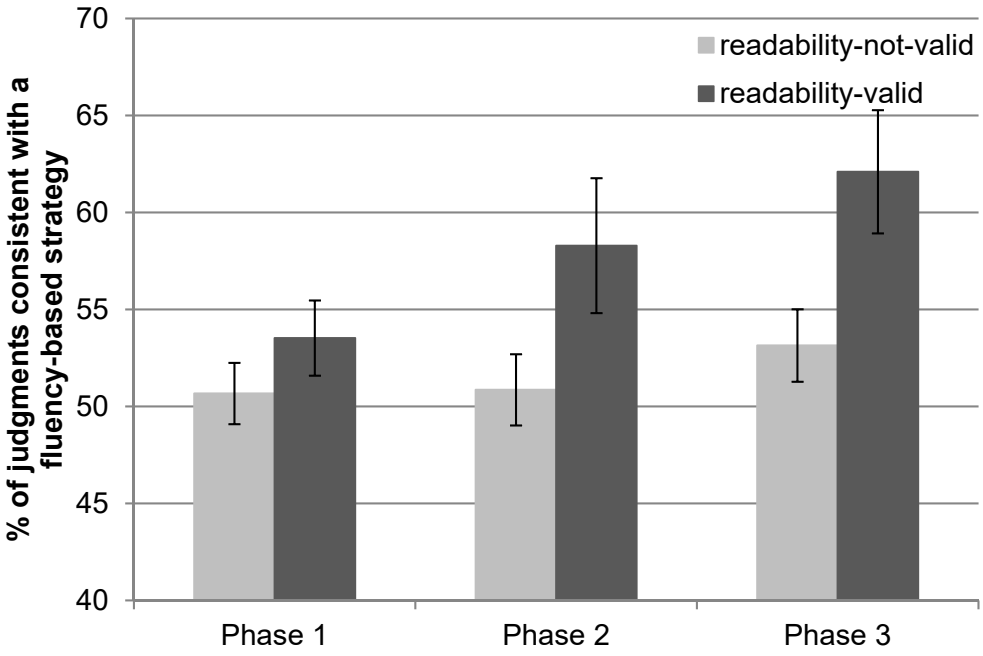

70
65
60
55
50
45
40
% of judgments consistent with a fluency-based strategy
readability-not-valid
readability-valid
Phase 1
Phase 2
Phase 3

Figure 3

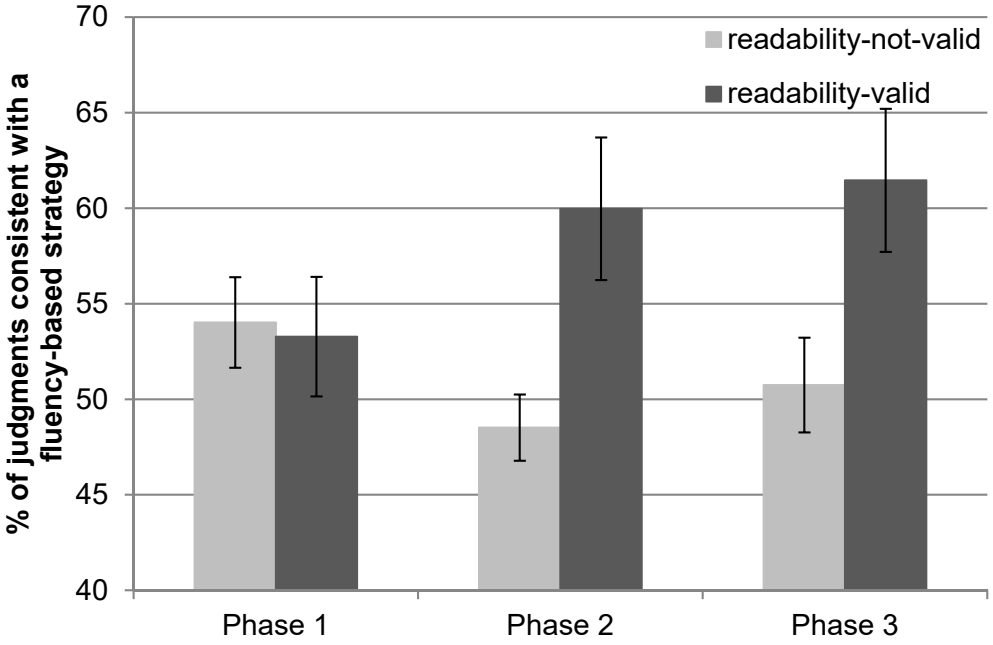

70
65
60
55
50
45
40
% of judgments consistent with a fluency-based strategy
readability-not-valid
readability-valid
Phase 1
Phase 2
Phase 3

Figure 4

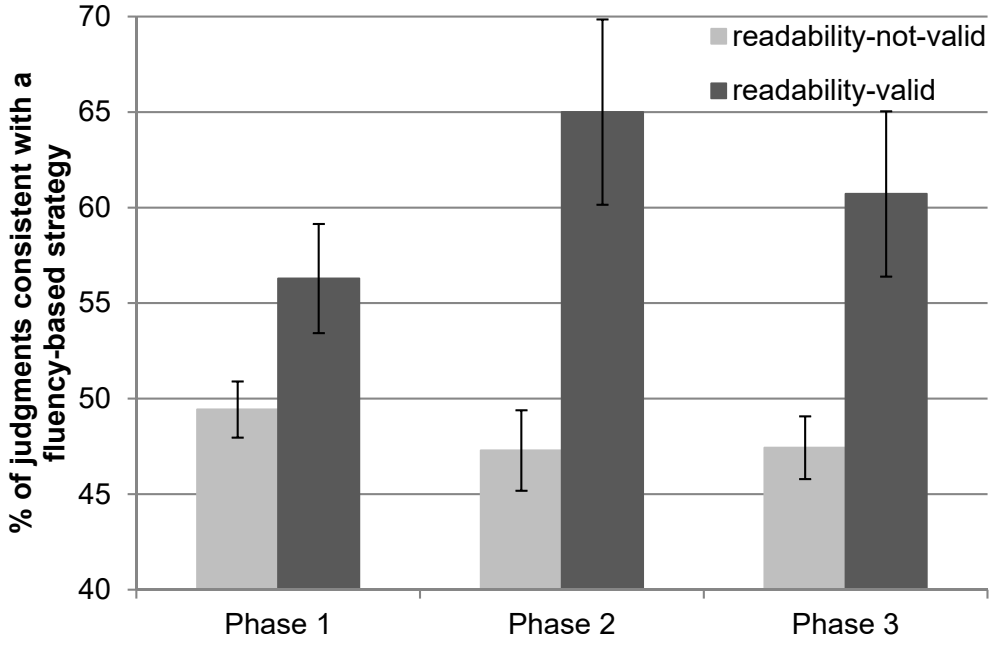
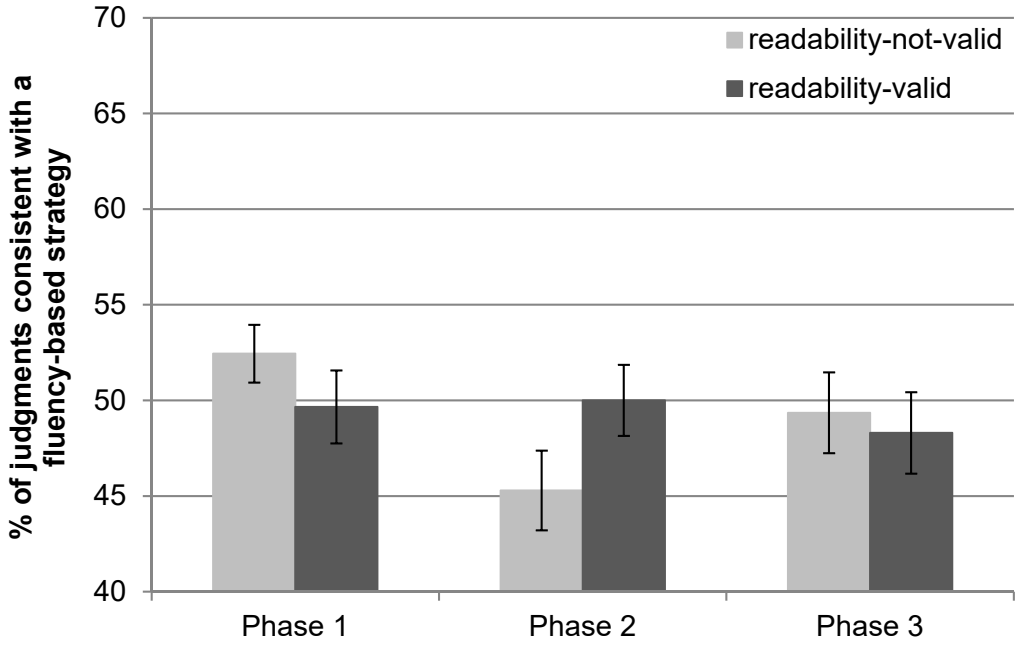

70
65
60
55
50
45
40
% of judgments consistent with a fluency-based strategy
readability-not-valid
readability-valid
Phase 1
Phase 2
Phase 3

Figure 5


70
65
60
55
50
45
40
% of judgments consistent with a fluency-based strategy
readability-not-valid
readability-valid
Phase 1
Phase 2
Phase 3

Footnotes

[1]For signal detection analysis, participants' responses to high fluent and low fluent statements are classified as hits, misses, correct rejections, or false alarms, with respect to the criterion truth status in the real world. Signal detection analysis allows for estimating the ability to discriminate between true and false statements (discrimination ability d'), and the tendency to respond „true" (response bias β; see Stanislaw & Todorov, 1999). In fluency-truth research, a significant difference in the response bias β for high fluent versus low fluent statements reflects a tendency to respond true to fluent, and false to less fluent statements, and is therefore indicative of a truth effect (Unkelbach, 2007).

[2]Because the color contrast manipulation has proven successful in prior research (e.g., Reber & Schwarz, 1999), it appears safe to assume that differences in perceptual fluency are elicited. Nevertheless, to support the claim that differences in fluency are elicited, we recorded latencies that reflect the interval between stimulus onset and decision. Under the assumption that it is stimulus perception but not the process of response formation that is affected by the color contrast manipulation, these latencies offer some indication about the manipulations' success (see Reber, Wurtz, & Zimmermann, 2004; Wurtz, Reber, & Zimmermann, 2008, for a discussion). For every experiment, these manipulation checks suggest that differences in perceptual fluency were elicited. Due to space limitations, full results are not reported in text, but are available from the authors.

[3]To address learning over trials, we fitted the percentage of judgments per trial in accordance with a fluency-based judgmental strategy (high readability=true; low readability=false) to a linear function separately for the readability-valid- and readability-not-valid conditions. Over trials, reliance on fluency was described by a linear function in the readability-valid-condition ($y = 47.48 + 0.24 * x$; $R^2 = .092$), $F(1,54) = 5.48$, $p < .03$, but not

in the readability-not-valid-condition (y = 45.48 + 0.06 * x; R^2 = .005), $F(1,54) = 0.27$, $p = .61$.

[4]Participants in the readability-valid condition indicated that 66.8% of the high readable statements and 50.8% of the low readable statements were true, while participants in the readable-not-valid condition indicated that 59.2% of the high readable statements and 55.5% of the low readable statements were true. Similar patterns were obtained in Experiments 3 and 4. Results are available from the authors.

[5]We thank one of the reviewers for suggesting the design of Experiment 5.

[6]At the end of each of the Experiments 1 to 4, participants were asked about the purpose of the present experiment. These open statements are some first helpful indicators with respect to participants' suspiciousness having received false feedback, conscious usage of a "high readability/fluency=true, low readability/fluency=false-strategy"strategy, and identification and application of a pattern matching strategy. Deception: None of the participants mentioned suspiciousness with respect to the feedback's validity. Conscious strategy: Five participants mentioned the influence of color-contrast, three participants the influence of readability, and four participants the influence of the appearance of the statements on decision making or truth judgments, however, without specifying or even addressing the fluency-truth correlation. Another participant addressed the influence of readability on response times (but not on truth). Finally, one participant—however, in a readability-not-valid condition—mentioned a positive readability-truth-correlation. Note that none of the participants mentioned "fluency" or the transfer from Phases 1 and 2 to Phase 3. Pattern matching: Only one participant mentioned that she had searched for but couldn't identify any patterns (despite the fact that she was in a readability-valid-condition), two other participants mentioned that the aim of the researchers was detecting patterns in decision

making but it remains a question of interpretation whether they also refer to their own strategies.

Because excluding these participants did not change the overall pattern of results in any of the experiments, we conducted all analyses in the manuscript with all participants.