Cognition 128 (2013) 252–255

Contents lists available at SciVerse ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Brief article

Disfluency prompts analytic thinking—But not always greater accuracy: Response to Thompson et al. (2013)

Adam L. Alter [a,*], Daniel M. Oppenheimer [b], Nicholas Epley [c]

[a] Stern School of Business, New York University, 40 West 4th Street, Suite 811, New York, NY 10012, United States
[b] Associate Professor of Marketing, UCLA Anderson School of Management, United States
[c] John T. Keller Professor of Behavioral Sciences, University of Chicago, Booth School of Business, United States

ARTICLE INFO

Article history:
Received 29 October 2012
Revised 7 January 2013
Accepted 9 January 2013
Available online 16 February 2013

Keywords:
Fluency
Disfluency
Dual-system processing
Metacognition
Heuristics
Analytical thinking
System 1 processing
System 2 processing

ABSTRACT

In this issue of *Cognition*, Thompson and her colleagues challenge the results from a paper we published several years ago (Alter, Oppenheimer, Epley, & Eyre, 2007). That paper demonstrated that metacognitive difficulty or disfluency can trigger more analytical thinking as measured by accuracy on several reasoning tasks. In their experiments, Thompson et al. find evidence that people process information more deeply—but not necessarily more accurately—when they experience disfluency. These results are consistent with our original theorizing, but the authors misinterpret it as counter-evidence because they suggest that accuracy (and even confidence) is a measure of deeper processing rather than a contingent outcome of such processing. We further suggest that Thompson et al. err when they discriminate between "perceptual fluency" and "answer fluency," the former of which is an element of the latter. Thompson et al. advance research by adding reaction time as a measure of deeper cognitive processing, but we caution against misinterpreting the meaning of accuracy.

1. Introduction

This issue of *Cognition* features a paper by Thompson et al. (2013) that both complements and challenges a paper we published several years ago (Alter, Oppenheimer, Epley, & Eyre, 2007). We're pleased that Thompson et al. engaged with our work, but also eager to address two areas of disagreement. First, Thompson et al. suggest that their findings fail to replicate ours. In fact, they confirm our account except when the questions are too difficult for participants to answer correctly. Both papers ultimately arrive at the same conclusion. Second, Thompson et al. discriminate between two forms of fluency—"perceptual fluency" and "answer fluency"—that we think are no more independent than apples are from fruit. Instead, as the analogy suggests, perceptual fluency is part of the ease with which an answer comes to mind. Since perceptual fluency corresponds to a genuine cognitive process, whereas answer fluency incorporates a range of operations, we question the value of comparing their relative contributions to judgment.

Our paper (Alter et al., 2007) consisted of four studies in which participants were less confident, and consequently more thoughtful, about their responses to various questions when they experienced disfluency—the subjective experience of difficulty while processing information. In one experiment, for example, participants were more likely to reconsider intuitively appealing but incorrect responses to three problems (the Cognitive Reflection Test, or CRT: Frederick, 2005) when the questions were printed in a difficult-to-read font. We theorized that participants misattributed the difficulty of reading the material for the difficulty of the problems themselves, thereby encouraging them to think more carefully. This careful thinking then improved performance on analytical reasoning tasks.

This basic effect of disfluency on more analytical processing has been replicated, extended, and applied a

* Corresponding author. Tel.: +1 212 998 0142.
E-mail address: aalter@stern.nyu.edu (A.L. Alter).

0010-0277/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.cognition.2013.01.006

number of times since our original demonstration (e.g., Cokely, Parpart, & Schooler, 2009; Diemand-Yauman, Oppenheimer, & Vaughan, 2011; Kappes & Alter, 2013; Oppenheimer & Alter, in press; Song & Schwarz, 2008; also see Simmons & Nelson, 2006, Study 13, for a relevant precursor). Like any scientists, we are pleased that Thompson et al. found the results important enough to examine further. Their approach is particularly appealing because it includes reaction times that serve as a more direct measure for analytical reasoning. Although other researchers have suggested that participants process information more analytically when they experience disfluency, Thompson et al. provide the first direct demonstration, to our knowledge, that disfluency prompts people to spend more time considering their responses. We agree with Thompson et al.'s suggestion that researchers should measure response times as well as accuracy rates to capture processing depth, and we believe this insight is the paper's primary contribution.

An alternative experimental approach (that we adopted in our experiments) is to choose tasks that inspire categorically different responses depending on whether participants are relying on intuitive and analytical thinking. This approach measures the depth of cognitive processing indirectly, depending on which category of response a participant exhibits. To be a valid measure of cognitive processing, however, this approach requires that the different types of cognitive processing actually lead to categorically different responses. If the items are so difficult that no amount of careful thinking will lead to a correct (or different) response, then a person's answer cannot tell us about her style of thinking.

Correct responses are therefore an inappropriate proxy for deeper processing when the reasoning tasks are either too difficult (and impossible to answer no matter much a person thinks about them), or too easy (and capable of being answered correctly based on intuition alone; for more on this point see Alter et al., 2007, Experiment 4). We believe that using overly difficult tasks to measure cognitive processing is precisely the reason for the apparent discrepancy between Thompson et al.'s results and our original results, and their paper consequently highlights the importance of calibrating cognitive tasks before using them to assess differences in processing depth.

2. Reconciling Thompson et al. (2013) and Alter et al. (2007)

Despite the differences in confidence and accuracy in Thompson et al. (2013) and Alter et al. (2007), both papers demonstrate the more fundamental principle that people process information more deeply when they experience cognitive disfluency. The differences arise because Thompson et al. find only limited evidence that disfluency improves accuracy. However, disfluency can encourage deeper processing without improving accuracy rates when the task is simply too difficult for participants to answer. Our theory proposed that disfluency triggers deeper cognitive processing, but deeper cognitive processing does not ensure accuracy in and of itself. Ten-year-olds are unable to solve calculus problems no matter how hard they think, and would be unable to solve them even if they are told directly that the problems are difficult and require considerable mental effort. The participants in Thompson et al.'s studies are certainly more sophisticated than ten-year-olds, but in both cases the questions are too difficult to uncover a link between processing depth and accuracy. The relationship between mental effort and mental performance is therefore similar to the relationship between physical effort and physical performance. None of the authors on this comment, for instance, can dunk a basketball no matter how hard he tries.

Thompson et al.'s results demonstrate that participants indeed thought more carefully about the tasks at hand, but were not always capable of providing the right answer even with more careful thought. In Experiments 1b, 2a, and 2b, participants spent more time generating an answer when they adopted a disfluent facial expression (Experiment 1b: $F[1, 70] = 8.27$, $p = .005$, $\eta_p^2 = .11$), or when the problems were presented in a disfluent font (Experiment 2a: $F[1, 46] = 10.86$, $p = .002$, $\eta_p^2 = .06$; Experiment 2b: $F[1, 40] = 4.07$, $p = .05$, $\eta_p^2 = .09$). As our theory would predict, disfluency triggers more careful and analytical thinking.

So why, if people thought more deeply, were their answers no more accurate than the answers of their peers in the studies' fluent conditions? The answer looks to be the same reason that none of us can dunk a basketball. Participants found the tasks to be too difficult even when they were explicitly instructed to think more carefully. For example, in Experiment 1a participants completed a series of syllogistic reasoning problems, first providing an intuitive response, and then thinking more carefully before providing a more deeply considered follow-up response. All responses were binary, requiring participants to indicate whether or not a conclusion followed logically from a preceding statement. Participants who responded randomly would therefore have answered approximately 50% of the questions correctly. As Thompson et al. reported in Table 1 of their manuscript, participants answered between 53% and 54% of the questions correctly when they relied on their intuitions, and only 55% of the questions correctly when they reconsidered their initial responses and thought about the problems more deeply. These overall accuracy rates are barely better than one would expect by chance alone, suggesting that the tasks were simply too difficult for additional thinking to improve performance. Indeed, the clearest way to identify a task as "difficult" is to show that overall accuracy rates are no better than chance.

In contrast to the results in Experiment 1a in Thompson et al., participants in Experiment 4 of Alter et al. chose their responses from among nine options, and managed to answer an average of 54% of the questions correctly across the fluent and disfluent conditions. These accuracy rates were therefore 43% higher than the 11% rate expected from chance alone, suggesting that their responses reflected a degree of skill. Because even a direct instruction to think harder failed to improve accuracy rates in Thompson et al.'s Experiment 1a, the questions were simply too difficult to serve the purpose for which they were chosen: to distinguish between shallower and deeper processing in the tested sample of students. Likewise, participants in