Journal of Experimental Social Psychology 49 (2013) 178–182

Contents lists available at SciVerse ScienceDirect

Journal of Experimental Social Psychology

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

FlashReport

Disfluency disrupts the confirmation bias

Ivan Hernandez *, Jesse Lee Preston

University of Illinois at Urbana-Champaign, USA

HIGHLIGHTS

- Confirmation bias is reduced when information is presented in a disfluent format.
- Less confirmation bias when evaluating capital punishment arguments (Study 1).
- Jurors give less confirmatory verdicts after reading a disfluent summary of a crime (Study 2).

ARTICLE INFO

Article history:
Received 28 June 2012
Revised 15 August 2012
Available online 5 September 2012

Keywords:
Fluency
Confirmation bias
Attitude change
Persuasion

ABSTRACT

One difficulty in persuasion is overcoming the confirmation bias, where people selectively seek evidence that is consistent with their prior beliefs and expectations. This biased search for information allows people to analyze new information in an efficient, but shallow way. The present research discusses how experienced difficultly in processing (disfluency) can reduce the confirmation bias by promoting careful, analytic processing. In two studies, participants with prior attitudes on an issue became less extreme after reading an argument on the issues in a disfluent format. The change occurred for both naturally occurring attitudes (i.e. political ideology) and experimentally assigned attitudes (i.e. positivity toward a court defendant). Importantly, disfluency did not reduce confirmation biases when participants were under cognitive load, suggesting that cognitive resources are necessary to overcome these biases. Overall, these results suggest that changing the style of an argument's presentation can lead to attitude change by promoting more comprehensive consideration of opposing views.

Introduction

People hold onto their beliefs strongly. Changing beliefs takes time and effort, and it is often easier to disregard alternative perspectives rather than to adapt existing beliefs. This preference for existing beliefs underlies the *confirmation bias* (Wason, 1960), where people seek out and interpret information that is consistent with their expectations (Koriat, Lichtenstein, & Fischhoff, 1980; Nickerson, 1998). Past reviews on the confirmation bias have conceptualized it (at least in part) as a cognitive shortcut or heuristic that simplifies complex inferential tasks (Friedrich, 1993; MacCoun, 1998). Because people assume that their existing beliefs are true, those beliefs serve as a heuristic for evaluating new information. However, the confirmation bias may also lead to poorer decisions because evidence is not being considered fully. For example, when children in a video were given a label of high or low socio-economic status, people used that label to make judgments of future academic ability, disregarding other relevant information (Darley & Gross, 1983).

The present research investigates how a seemingly irrelevant feature of a message—*fluency*—can lead people to re-evaluate information on previously formed attitudes and reduce confirmation bias effects. *Fluency* may be defined as the relative ease experienced during processing, and can be altered by features such as the visual clarity of text. Research shows that the effort associated with disfluency prompts a deeper, more analytical, and critical processing of the information itself. For example, high school students who received classroom materials in a disfluent font (e.g. Comic Sans italicized) scored higher on their examinations than when the material was presented fluently (Diemand-Yauman, Oppenheimer, & Vaughan, 2011), because the increased effort during reading led to processing the material more comprehensively. Similarly, people are more likely to give correct answers to "trick" questions that require reflection and the rejection of an intuitive response (e.g., Song & Schwarz, 2008; Alter, Oppenheimer, Epley, & Eyre, 2007, Study 1), and also evaluate consumer products based on systematic cues (such as quality) rather than heuristic cues (such as attractiveness) when presented disfluently (Alter et al., 2007, Study 2). By promoting metacognitive difficulty, disfluency can lead to deeper processing of the information and less reliance on heuristic modes of processing. In contrast, the ease that is facilitated by fluency leads to greater use of heuristic reasoning, and sometimes more errors in judgment.

Present research

The present studies examine whether the experience of disfluency can decrease confirmation bias effects when making evaluations of

* Corresponding author at: University of Illinois at Urbana-Champaign, 603 East Daniel Street, Champaign, IL 61820, USA.
E-mail address: Hernan27@illinois.edu (I. Hernandez).

0022-1031/$ – see front matter
http://dx.doi.org/10.1016/j.jesp.2012.08.010

information related to pre-existing attitudes. The increased difficulty experienced in processing disfluent text should make people more critical and analytical of that information, allowing for greater consideration of counter-attitudinal arguments, and more skepticism toward attitude-consistent information. As discussed, disfluent presentation has been shown to impact the effortful processing and evaluation of new information. But to date, no work has examined whether disfluency can also affect the processing of information where a person has already formed attitudes, as in the confirmation bias. Two studies tested this prediction, with pre-existing political attitudes (Study 1) and experimentally manipulated attitudes (Study 2).

Study 1: ideological disconfirmation

Previous studies on the confirmation bias have found that participants with strong prior beliefs on social issues like capital punishment evaluate related information in a manner that is consistent with their prior beliefs (Lord, Ross, & Lepper, 1979). Study 1 examined whether the confirmation bias may be reduced when information is presented disfluently. Participants read pro-capital arguments presented in either a fluent or disfluent font before making evaluations of the reading. Because conservatives show greater support for capital punishment than liberals (Carroll, 2004), we expected that people would make judgments consistent with their political ideology when the arguments were presented in a fluent font, but the bias would decline when information is presented in a disfluent font.

Method

Participants

133 undergraduates, (60 women, 73 men, $M_{age} = 19$) volunteered to participate in exchange for partial course credit.

Procedure

Instructions and stimuli. Participants were randomly assigned to a fluent or disfluent condition at the beginning of the study. All instructions and measures were completed on a computer program, in a private lab room. Participants first completed a short demographics questionnaire, which included questions about age, gender, ethnicity, and religious affiliation. Participants reported their political ideology on a 7-point scale, with endpoints 1 = "strongly liberal", 7 = "strongly conservative."

Next, participants read a short article in favor of capital punishment, used in previous research (Blanchard-Fields & Horhota, 2005). In the fluent condition, the article was presented in a 12-point Times New Roman font. In the disfluent condition, the article was presented in a light gray bold and italicized Haettenschweiler font, used in previous research to induce processing difficulty (e.g., Diemand-Yauman et al., 2011). After reading, participants answered six questions on the article. Three items related to assessments of the author — i.e., how *considerate*, and *understanding* the author seemed, and how *mature* the writing seemed. Three items measured agreement with the article, our key variable of interest. Specifically, participants were asked "How reliable is the message?", "How intelligent do you consider the argument?" and "How much do you believe the facts that were in the reading?" All questions were asked individually on 5-point Likert scales with endpoints 1 = "not at all", and 5 = "extremely".

Results

The three dependent measures (message reliability, intelligence of argument, belief in the argument) showed good scale reliability, Cronbach's $\alpha = .77$, and so the average of the items was computed to create a single composite measure of agreement. Our prediction that disfluency would reduce the confirmation bias effect of ideology was analyzed using a multiple regression. Political ideology was standardized with a mean of 0 and a standard deviation of 1, such that higher values indicated greater conservatism. Next, fluency condition was dummy coded with participants who saw fluent materials scored as a 1, and participants who saw the disfluent article scored as a 0. The product of these two variables was calculated to create the interaction term for the multiple regression. We entered the variables of fluency, standardized ideology, and the interaction term into a multiple regression to predict the composite measure of agreement. There were no significant main effects for ideology ($\beta = .04$, *ns*) or for fluency ($\beta = .02$, *ns*). However, the lack of main effects was qualified by the predicted significant interaction between expectation and fluency ($\beta = .26$, $p < .05$). The positive coefficient implies that fluency was associated with a partisan bias toward the arguments (i.e. that conservative participants agreed more with the arguments than liberals), and disfluency was associated with less partisan agreement (see Fig. 1).

To test whether participants used ideological biases *only* in the fluent condition, we conducted two separate regressions of conservatism on agreement for the fluent and disfluent conditions. The effect of ideology on agreement for participants who saw the *fluent* arguments is significant ($R^2 = .13$, $F(1,70) = 10.34$, $p < .01$), indicating that people's bias affected their judgments for fluent materials. However, the relationship between ideology and agreement for the *disfluent* arguments was not significant ($R^2 = .001$, $F(1,59) = .09$, *ns*) indicating that participants' prior ideological bias did not affect their subsequent judgment of the capital punishment arguments. These results support our hypothesis where compared to the partisan judgment in the fluent condition, when the arguments are presented in a disfluent format, conservatives and liberals are less persuaded by prior attitudes and responded more moderately.

Discussion

These results provide evidence that disfluency can attenuate ideological bias on judgment of political issues. When a political issue was presented fluently, conservatives and liberals were highly polarized in their judgment, consistent with their prior attitudes. But this confirmatory pattern was reduced when the arguments were presented disfluently, suggesting that the disfluency prompted more careful critical analysis of the argument.

However, there are two important limitations to Study 1. Because the biases observed in Study 1 were pre-formed attitudes, and not randomly assigned to the participants, it is possible that the observed effect was the result of some self-selective factor (e.g. people who naturally

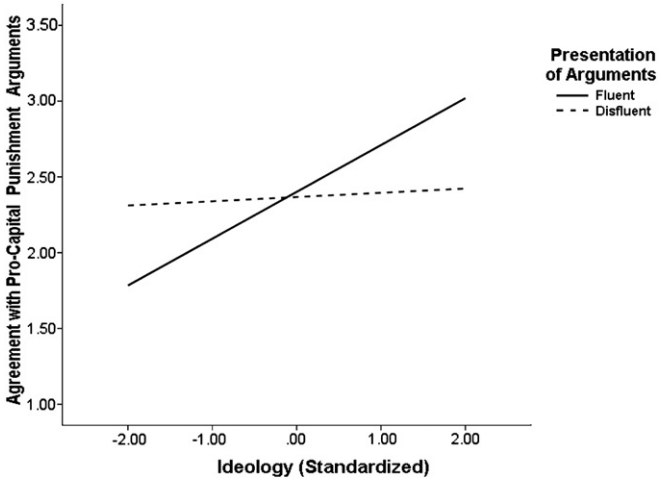

Fig. 1. Fitted values for mean agreement with pro-capital punishment arguments from regression of fluency of argument (1 = fluent, 0 = disfluent) and standardized measure of political ideology, with higher values denoting more conservatism (Study 1).

form strong attitudes). In addition, it is also possible that the disconfirmation effect observed in the disfluent conditions was due to disengagement or ambivalence toward the material, resulting in a general regression toward moderate responses. Our second study sought to address these limitations by experimentally manipulating prior bias, and by examining the role of attention and cognitive resources.

Study 2: legal judgments

One domain used extensively in confirmation bias research is juror decision making. This research finds that the final verdicts decided by juries usually conform to the tentative decisions they initially formed (Lawson, 1968) and this confirmation bias may be due to fast processing that only considers a subset of information (Pennington & Hastie, 1993). Study 2 examined whether processing difficulty could potentially lead to less biased verdicts in a mock-trial, in an experimental setting. Additionally, we hope to demonstrate that this disconfirmation is due to deeper analytic processing that requires attention and cognitive resources, rather than general disengagement and apathy in the disfluent condition. Participants played the role of a juror where they were given either a positive or negative description of a defendant and then read a description of the alleged crime in either a clear (fluent) or degraded (disfluent) format. In addition, we included two conditions that combined a disfluent text with a cognitive load manipulation (i.e. time constraint, memorization task). As in Study 1, we predicted a confirmation bias in the fluent condition, but not in the disfluent condition. Consistent with an analytical processing explanation, we also predicted that disfluency would not reduce confirmation biases when participants are under cognitive load. If disconfirmation is the result of more critical analysis in disfluent conditions, the effect should only be observed when cognitive resources are available. However, if the effects are the result of disengagement with the disfluent text, cognitive load should not interfere with the disconfirmation pattern observed in disfluent conditions.

Method

Participants

408 participants (144 women, 259 men, and 5 no response; M_{age} = 27) completed a brief questionnaire on Amazon.com's Mechanical Turk survey website. IP addresses were recorded to ensure unique responses, and 7 participants were removed due to double participation, and 3 for having no response, for a final sample size of 398.

Procedure

Participants were randomly assigned to a condition in a 2 (bias: positive/negative) × 4 (condition: fluent/disfluent/time constraint/memory load) between-subjects design. All participants were instructed that they would read information about a defendant accused of a crime, and would be asked to decide on the verdict. Following the instructions, participants in the time constraint condition were presented a clock timer, and were told to submit their responses as soon as the timer reaches 3 min. In the memorization condition, participants were briefly shown a list of words (guitar, eagle, glasses, mixer, ocean, table, parade, window, baseball), and asked to keep them in memory throughout for later recall.

Next, participants read a witness testimony from the defendant's school psychologist about her interactions with the defendant, Donald. In the positive bias condition, Donald was described as having a history of great remarks from teachers, polite, warm, good sense of humor, respectful, easy to work with, and a good listener. In the negative bias condition, Donald was described as having a history of disciplinary issues, rude, cold, criticizing, disrespectful, difficult to work with, and interrupting.

Participants then read a description of the objective facts of the case, where Donald was accused of a robbery, but his guilt is ambiguous. In the fluent condition, the facts were written in a Times New Roman 16-point font. In the disfluent conditions, participants received a document written in a 12-point Times New Roman font that had been photocopied recursively three times on the lowest contrast setting until the text was significantly degraded, but still readable, which has been shown to induce analytic thinking via disfluency (Diemand-Yauman et al., 2011; Oppenheimer & Frank, 2008).

After reading the documents, participants decided their verdict (1 = guilty/0 = not guilty) and length of jail sentence (0–5 + months). Additionally, participants were asked to rate the certainty of their verdict on a seven point scale (endpoints: 1 = extremely certain he is not guilty, 7 = extremely certain he is guilty) and how interested they were in the study (endpoints: 1 = extremely bored, 7 = extremely interested).

Results

First, we compared the means of guilty verdicts between the two cognitive load conditions with a 2 (bias: positive/negative) × 2 (condition: time constraint/memory load) ANOVA. There was an expected main effect for bias $F(1,165) = 16.22$, $p < .001$, but no significant main effect between the two conditions ($F(1,165) = .09$, $p = .77$) or two-way interaction ($F(1,165) = 2.48$, $p = .12$) (see all means in Table 1). We therefore collapsed these together into a single "cognitive load" condition for subsequent analyses. Our main dependent variables were each analyzed with a 2 (bias: positive/negative) × 3 (condition: fluent/disfluent/cognitive load) ANOVA. On the dichotomous measure of verdict (0 = not guilty, 1 = guilty), there was no main effect for condition ($F < 1$), but the main effect of bias was significant ($F(1,392) = 17.11$, $p < .001$), with more guilty verdicts in the negative ($M = .73$; $SD = .45$) compared to the positive bias condition ($M = .51$; $SD = .50$). Most importantly, the predicted interaction between bias and condition was significant, $F(2,392) = 3.51$, $p < .05$. We conducted simple effects tests to test the specific hypothesis that confirmation bias effects would be observed in the fluent and cognitive load conditions, but not in the disfluent condition. As predicted, participants who read the fluent information were more lenient when they had a positive vs. negative bias toward Donald, $F(1,392) = 12.26$, $p < .01$. Likewise, participants under cognitive load also showed an effect of bias, $F(1,392) = 15.11$, $p < .001$. Of key importance, however, there was no effect of bias in the disfluent condition, $F(1,392) = .02$, $p = .90$, with moderate judgments of guilt in both positive and negative bias conditions, (see all means in Table 1).

Next, we analyzed the continuous measure of guilt certainty by a 2 (bias) × 3 (condition) ANOVA. There was no main effect for condition ($F(1,392) = 1.78$, $p = .17$), and a main effect of bias ($F(1,392) = 26.51$, $p < .001$). Most importantly, the predicted interaction between bias

Table 1
Percentage of guilty verdicts given to defendant, by Fluency of crime description condition and prior impression of the defendant (Study 2).

	Reading condition			
	Fluent	Disfluent	Disfluent + Time constraint	Disfluent + Memory Load
Positive bias	48%	58%	55%	41%
Negative bias	78%	60%	72%	81%

and condition was significant ($F(1,392) = 3.62$, $p < .05$). Simple effects tests reveal that bias had a significant effect in both the fluent condition, ($M_{positive} = 3.83$, $SD = 1.77$; $M_{negative} = 4.91$, $SD = 1.72$; $F(1,392) = 11.60$, $p < .01$), and in the cognitive load condition ($M_{positive} = 3.68$, $SD = 1.72$; $M_{negative} = 5.05$, $SD = 1.70$; $F(1,392) = 27.08$, $p < .001$). But as predicted, there was no simple effect of bias in the disfluent condition ($M_{positive} = 4.60$, $SD = 1.78$; $M_{negative} = 4.86$, $SD = 1.60$, $F < 1$).

Length of jail sentence was also analyzed by a 2 (bias) × 3 (condition) ANOVA. There was a main effect of bias $F(1,389) = 55.00$, $p < .001$, with longer sentences given in the negative ($M = 2.62$, $SD = 1.86$) vs. the positive bias condition ($M = 1.25$, $SD = 1.61$). Neither the main effect of condition ($F < 1$) nor the two-way interaction was significant, ($F(2,389) = 1.01$ $p = .36$). The null interaction was not predicted, as we expected the length of jail sentences to follow the same pattern as guilty verdicts. One possible reason for this discrepancy is that the length of jail sentences ranged from zero to five months or more, but the full range of the scale was not used by participants who judged Donald as not guilty. These participants' scores were anchored at the low end (0 months), whereas participants who judged Donald as guilty used the full range of the scale from 0 to 5. As a result, there was a robust effect of bias on length of jail sentence, but the interaction had relatively low statistical power.

Finally, we also conducted a 2 × 3 ANOVA on ratings of interest in the study. There were no main effects for bias or condition, $Fs < 1$, nor an interaction $F(2,392) = 2.43$, $p = .09$.

Discussion

In Study 2, participants in a mock-trial study showed confirmation bias effects when information was presented fluently, but not disfluently. Importantly, confirmation biases were also observed when participants read disfluent information under cognitive load (time pressure, memory load). The key role of cognitive resources helps rule out an alternative disengagement account of disfluency, i.e. that moderate responses are explained by regression to the mean. Because the disconfirmation effect was not observed under cognitive load such as distraction or time pressure, it is unlikely these effects are due to disengagement with the material. However, while these effects are suggestive, we acknowledge that this evidence does not directly demonstrate that the effect is due to deeper processing. It is possible that some other aspect of disfluency interfered with the confirmation bias. But these results do indicate that cognitive resources are *necessary* for disconfirmation to occur. Any other explanation would also have to account for the confirmation bias observed in the disfluent/cognitive load conditions. We suggest that these results are most consistent with an analytical processing explanation, consistent with previous research.

General discussion

In two studies, decreasing the processing ease of an argument reduced confirmation bias effects, the first evidence that disfluency can affect processing of information where one has pre-existing attitudes. In Study 1, both conservatives and liberals evaluated a capital punishment article consistent with their political beliefs when it was presented fluently, but the confirmation bias declined when the argument was presented disfluently. Study 2 replicated this effect with assessments of guilt, where participants gave less biased judgments when the facts were presented disfluently. Most importantly, Study 2 found that people were only able to disconfirm their prior biases when they had cognitive resources available, consistent with other evidence that disfluency can prompt an analytical and critical mindset (e.g., Oppenheimer, 2008). Because participants who are mentally occupied do not demonstrate disconfirmation, we can rule out a disengagement explanation. Previous work has found that disfluency can impact abstract thinking (Alter & Oppenheimer, 2008; Yang, Preston, & Hernandez, in press), and improve reflective reasoning skills (Song & Schwarz, 2008), and these findings may show another area where people benefit from the effect of disfluency on analytic thinking. Just like speed bumps cause one to drive more slowly and carefully, the experience of difficulty associated with disfluency prompts a slower, more careful mindset when making judgments, even when one comes to the issue with existing biases.

Our results also highlight the importance of expectations on the observed effects of fluency. While the majority of research examines the main effect of fluency on judgment, recent research supports an interactional model that includes the *context* of the fluency/disfluency (Oppenheimer, 2008). For example, people rate fluent essays more highly only when they are told that fluency is a sign of positivity (i.e. implies intelligence), and more negatively when told that fluency is negative (i.e. simple-mindedness) (Brinol, Petty, & Tormala, 2006). Similarly, the effect of fluency is moderated by people's goals. People gave fluent targets more positive ratings, but only in domains where they felt fluency was a desirable quality (e.g. when reading for pleasure). When people had a goal that implied difficulty (e.g. acquiring information), disfluent texts were rated more favorably (Galak & Nelson, 2011). Our present results likewise demonstrate that fluency does not produce uniformly positive or negative evaluations of a target, rather the direction of attitude change depends on how disfluency is interpreted in the context of expectations and existing attitudes.

In sum, these findings suggest a simple and promising tool for persuasion and for overcoming biases that can often distort reasoning. Disfluency may offer an opportunity for better judgment and discourse between opposing positions, ultimately giving what was once an overlooked message, a chance to be seen.

Acknowledgments

We thank How Hwee Ong, Joy Menhennett, Suzy Millman, and Ben Chesher for their assistance in data collection. We also thank Ryan Ritter, Mike Morrison, and Dov Cohen for helpful comments on a previous draft.

References

Alter, A. L., & Oppenheimer, D. M. (2008). Effects of fluency on psychological distance and mental construal (or why New York is a large city, but New York is a civilized jungle). *Psychological Science*, *19*, 161–167.
Alter, A. L., Oppenheimer, D. M., Epley, N., & Eyre, R. N. (2007). Overcoming intuition: Metacognitive difficulty activates analytic reasoning. *Journal of Experimental Psychology. General*, *136*, 569–576.
Blanchard-Fields, F., & Horhota, M. (2005). Age differences in the correspondence bias: When a plausible explanation matters. *Journals of Gerontology: Psychological Sciences*, *60*, 259–267.
Brinol, P., Petty, R. E., & Tormala, Z. L. (2006). The malleable meaning of subjective ease. *Psychological Science*, *17*, 200–206.
Carroll, J. (2004). *"Who supports the death penalty?" Gallup Poll News Service, November 16, 2004.* The Gallup Organization.
Darley, J. M., & Gross, P. H. (1983). A hypothesis confirming bias in labeling effects. *Journal of Personality and Social Psychology*, *44*, 20–33.
Diemand-Yauman, C., Oppenheimer, D. M., & Vaughan, E. B. (2011). Fortune favors the bold (and the italicized): Effects of disfluency on educational outcomes. *Cognition*, *118*, 114–118.
Friedrich, J. (1993). Primary error detection and minimization (PEDMIN) strategies in social cognition: A reinterpretation of confirmation bias phenomena. *Psychological Review*, *100*(298), 319.
Galak, J., & Nelson, L. D. (2011). The virtues of opaque prose: How lay beliefs about fluency influence perceptions of quality. *Journal of Experimental Social Psychology*, *47*, 250–253.
Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning & Memory*, *6*, 107–118.
Lawson, R. (1968). Order of presentation as a factor in jury persuasion. *Kentucky Law Journal*, *56*, 523–555.
Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, *37*, 2098–2109.
MacCoun, R. J. (1998). Biases in the interpretation and use of research results. *Annual Review of Psychology*, *100*, 298–319.
Nickerson, R. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*, 175–220.

Oppenheimer, D. M. (2008). The secret life of fluency. *Trends in Cognitive Science*, *12*, 237–241.
Oppenheimer, D. M., & Frank, M. C. (2008). A rose in any other font wouldn't smell as sweet: Effects of perceptual fluency on categorization. *Cognition*, *106*, 1178–1194.
Pennington, N., & Hastie, R. (1993). The story model for juror decision making. In R. Hastie (Ed.), *Inside the juror: The psychology of juror decision making* (pp. 192–221). New York, NY: Cambridge University Press.
Song, H., & Schwarz, N. (2008). Fluency and the detection of misleading questions: Low processing fluency attenuates the Moses illusion. *Social Cognition*, *26*, 791–799.
Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology*, *12*, 129–140.
Yang, D. Y-J., Preston, J. L., & Hernandez, I. (in press). Polarized attitudes toward the Ground Zero mosque are reduced by high-level construal. *Social Psychological and Personality Science*, http://dx.doi.org/10.1177/1948550612446973.