ARTICLE IN PRESS

Arguing against confirmation bias: The effect of argumentative discourse goals on the use of disconfirming evidence in written argument

Constanza Villarroel[a],*, Mark Felton[b], Merce Garcia-Mila[a]

[a] Departament of Cognition, Development and Educational Psychology, University of Barcelona, Spain
[b] Lurie College of Education, San José State University, San Jose, CA, USA

ARTICLE INFO

ABSTRACT

This study explores the impact of argumentative discourse goals on confirmation bias in young adults. All participants were presented three types of graphical evidence: data supporting their initial view, challenging their initial view and ambiguous data that could be interpreted either way. They were asked to use the evidence to write argumentative essays before and after engaging in a chat-based dialogue with a partner who held an opposing view. Dyads were assigned to one of two argumentative discourse goal conditions: Argue to persuade or argue to reach consensus. At the posttest, participants in the persuasion condition were more likely to misinterpret evidence and less likely to reference their dialogue than peers in the consensus condition. Educational implications are discussed.

1. Introduction

Confirmation bias, defined as the tendency to select and interpret evidence in ways that reinforce prior beliefs, is a widespread and well documented phenomenon (Klayman, 1995; Mynatt, Doherty, & Tweney, 1977; Nickerson, 1998). Klayman (1995) suggests that it is most likely a constellation of interrelated cognitive biases that are best understood in terms of how we look to confirm our beliefs on the one hand, and how we neglect to challenge our beliefs on the other. He argues that these complementary processes should be studied in tandem, since the tendency to seek confirming evidence, in itself, does not violate the norms of effective reasoning. Similarly, Nickerson (1998) suggests that the tendency to look for confirming evidence may be a natural part of inductive reasoning. However, the problem with the active search for confirming evidence is that it can leave individuals overconfident about their beliefs if they have failed to give full and fair consideration to evidence that challenges their beliefs.

Another facet of confirmation bias, aside from the biased search for evidence, is the biased evaluation of evidence (Ditto & Lopez, 1992; Taber & Lodge, 2012). This is the bias examined by Lord, Ross and Lepper (1989) in their seminal study on the effects of mixed evidence on attitude polarization. In this study, they exposed participants to two studies on capital punishment, one that supported and one that challenged its effectiveness as a deterrent to murder. The researchers found

that participants' opinions grew more polarized despite the fact that half of the evidence presented to them was inconsistent with their view on the topic. They suggest that in response to mixed evidence, participants were inclined to overvalue confirming evidence and undervalue disconfirming evidence, leading them to feel more certain of their original position. To look more directly at the question of how participants interpret mixed evidence, Klaczynski and Gordon (1996) presented adolescents with two studies, each with flaws to internal validity, which either supported or challenged participants' religious views. They found that participants were more likely to identify the threat to validity in the study that challenged their religious views, and less likely to identify the same threat in the study that supported their view. This suggests that confirmation bias may trigger a form of motivated skepticism of opposing evidence when reasoning. In addition, Taber and Lodge (2012) found that individuals spend more time processing disconfirming evidence than confirming arguments when reading texts about controversial issues on which they held strong beliefs (gun control and affirmative action). And when prompted to respond in writing to arguments taken from these texts, participants produced more critical reactions (emotional and cognitive) to opposing side arguments than either favorable or critical reactions to same-side arguments and evidence.

In short, the natural tendency to build theories from confirming evidence can create an internal monologue with an amplifying effect as individuals take in more confirming evidence and become more skeptical of disconfirming evidence (Ditto & Lopez, 1992; Jonas, Schulz-Hardt, Frey, & Thelen, 2001; Nisbett & Ross, 1980; Taber & Lodge, 2012). Working from this model, several researchers have investigated the effects of dialogic argument on confirmation bias, under the presumption that engaging with an opposing voice might disrupt this tendency towards monologue and one-sided reasoning (Kuhn, Shaw & Felton, 1997; Reznitskaya et al., 2001).

1.1. Dialogic argument and confirmation bias

According to some theorists, all argumentation – including argumentative reasoning – is fundamentally an exchange with a real or imagined audience (Billig, 1996; Walton, 1989). Taking this view, argumentative dialogue can have a profound impact on reasoning by providing opportunities for individuals to elaborate ideas, generate and interrogate evidence, and reflect on conclusions with a skeptical partner (Asterhan & Schwarz, 2009; Felton & Kuhn, 2001; Rips, Brem, & Bailenson, 1999). Argumentative dialogue, in turn, might offset confirmation bias by creating a context in which opposing sources of evidence are voiced and considered as the burden of proof shifts from one speaker to the other. If so, social engagement would produce opportunities for cognitive conflict around the interpretation and evaluation of data and promote more two-sided reasoning. Several studies have confirmed that argumentative dialogue can have this effect on individual argumentation (Dong, Anderson, Kim, & Li, 2002; Gillies & Khan, 2009; Kuhn, Shaw, & Felton, 1997; Reznitskaya et al., 2001), though the effect is most robust where the dialogue is structured. For example, Felton (2004) found that argumentative dialogue plus metacognitive reflection on dialogue increased two-sided argumentation in both acknowledging and addressing other-side claims. While these studies have not looked directly at how individuals select, interpret and present evidence in their writing, they do suggest that under the right conditions, dialogic argumentation can push individuals to address alternative claims when framing a written argument.

And yet, despite evidence that dialogic argument can mitigate confirmation bias, there is also evidence that dialogue can have the opposite effect. Kuhn and Lao (1996) found that, similar to mixed evidence, dialogue with an opposing partner increases polarization among individuals with strong initial opinions for or against capital punishment. They propose that individuals with more fixed, stable and integrated opinions may be more resistant to evidence that might otherwise lead them to moderate their views. In a meta-analysis of studies on confirmation bias (which they refer to as *congeniality bias*), Hart and colleagues (Hart et al., 2009) point out that cognitive conflict can produce negative affect (Harmon-Jones, 2000), which can then evoke a *defense motivation* in individuals. They found moderate support for the hypothesis that defense motivation is a moderating variable for confirmation bias, suggesting that biased assimilation of evidence and attitude polarization may occur when participants feel that their position is being threatened. In such situations, the motivation to defend a position may trigger the tendency to overvalue confirming evidence and undervalue or misinterpret disconfirming evidence.

Here it is important to note that Nickerson (1998) argues that confirmation bias should be understood as an unconscious reasoning process and distinguished from the conscious and strategic case-building individuals may employ to win an argument or defend a position. However the dividing line between conscious and unconscious processes may not be so clear, particularly in situations where individuals are motivated to preserve their prior beliefs. Research in my-side bias, the tendency to build one-sided arguments when presenting a position, suggests that the phenomenon may be driven by unexamined schema for good argument, which can be mitigated with explicit instruction (Nussbaum, Kardash, & Graham, 2005; Wolfe & Britt, 2008; Wolfe, Britt, & Butler, 2009). If confirmation bias operates under similar principles, decreasing defense motivation combined with increasing conscious awareness of alternative interpretations of disconfirming evidence may diminish its effect on reasoning and offer further insight into its root cause.

1.2. Argumentative goals and confirmation bias

Conflicting evidence regarding the effects of argumentative dialogue on confirmation bias suggests that more research is needed to unpack their relationship. On the one hand, argumentative dialogue may promote active re-evaluation of prior beliefs via cognitive conflict; on the other hand, it may trigger defense motivation and increase skepticism towards