OSSA Conference Archive   OSSA 10

May 22nd, 9:00 AM - May 25th, 5:00 PM

Know thy biases! Bringing argumentative virtues to the classroom

Frank Zenker

Lund University, Department of Philosophy and Cognitive Science

Know thy biases! Bringing argumentative virtues to the classroom

FRANK ZENKER

Department of Philosophy and Cognitive Science
Lund University
Kungshuset, Lundagård, 22 222 Lund
Sweden
frank.zenker@fil.lu.se

ABSTRACT: We present empirical evidence from social psychological research which suggests that standard methods ("show & tell") employed when teaching the heuristics and biases program in the context of critical thinking (CT) instruction are likelier to facilitate the discernment and correction of biases in others' reasoning than to have a similar effect in the self-monitoring case. Exemplified by the social phenomenon of false polarization, we suggest that CT instruction may be improved by fostering student's abilities at counterfactual meta-cognition, and present a corresponding teaching and learning activity.

KEYWORDS: biases, counterfactual meta-cognition, critical thinking instruction, false polarization, teaching and learning activity

1. INTRODUCTION

According to the heuristics and biases program in social psychology (e.g., Gigerenzer et al., 1999; Kahneman et al., 1982; Kahneman, 2011; Sahlin et al., 2010; Tversky & Kahneman, 1974), when acquiring and evaluating information in order to acquire, sustain, or change belief, humans regularly entertain search rules (aka heuristics) that are under-complex (*biased*) and regularly terminated before an optimal result is achieved (*satisficing*). It is natural to assume that making agents aware of such biases would also tend to de-bias them—i.e., minimize the effects of biases in their own reasoning and belief. However, experimental evidence suggests that merely teaching people that they are likely subject to biases is a rather ineffective means of de-biasing them. A "deeper" engagement with agents' cognitive skills and practices appears to be required in order for de-biasing to function as a self-corrective contribution to better judgment and action.

The following motivates and describes a teaching and learning activity (TLA) which should improve learners' ability to self-discern and self-correct their own biases. The purpose is to assist instructors of critical thinking courses in designing and deploying this TLA either in a dedicated course or as an instrument in "critical thinking across the curriculum." The TLA is based on experimental studies carried out in social psychology, but it has (to the author's knowledge) not been systematically tested in a class room setting, so far. In the terms of Fischhoff (2010), the TLA constitutes an item of *prescriptive interventionism*.

After briefly describing biases (Sect. 2), we present our test case, false polarization, as a social phenomenon that is sustained by biases and their ascription (3). Subsequently, we lay out a route along which the ability to self-discern biases may be improved (4). We then describe the TLA, both analytically and sequentially, formulate learning outcomes and assessment techniques (5), and close with a discussion (6) plus summary (7).

2. BIASES

Larrick (2008, p. 319f.) offers a useful distinction between three basic types of errors in human reasoning and decision making that are owed to biases: (i) *psychophysically-based errors* such as reference point effects, e.g., the overweighing of the most recently received information *vis-à-vis* all information received; (ii) *association-based errors*, e.g., deciding on the basis of so far available evidence rather than searching for more complete/representative evidence; and (iii) *strategy-based errors*, e.g., the use of inferior strategies or decision rules, such as "a bird in the hand is worth two in the bush." As he further points out, "many biases are multiply-determined" (ibid.)

Biases, and the framing effects to which they can give rise (see below), are considered *latent*: subjects are normally unaware that a bias is "at work." This being so may be found explained in terms of an evolutionary inheritance as the result of "coping" with natural and social environments in ways that deliver acceptable results in some or perhaps even most cases. It is therefore only half-joking to speak of "Simple heuristics that make us smart [enough to get by]" (Gigerenzer et al., 1999; Gigerenzer & Brighton, 2009).

With some simplification, what contemporary social psychologists have come to call "biases," philosophers and scholars of law, amongst others, have long known as "fallacies." The latter two fields share a tradition in Aristotelian scholarship, specifically the critiques of the Sophistic mode of audience persuasion—in response to which the identification and classification of fallacies arose as a scholarly topic (Conley, 1990). Amongst others, the 16th century Francis Bacon's *idolatry* or the 17th century John Locke's naming of a range of fallacies fronted by 'ad' (e.g., *ad hominem*) have continued this tradition into the modern age. With Hamblin's (1970) standard work, fallacies have come to be "bread and butter" in speech communication, rhetoric, and argumentation studies.

In the discussion that is nowadays centered in the cognitive sciences, widely understood, terms such as "anchoring" are—for reasons one might speculate about—often treated as recent discoveries, although predecessor terms may almost always be identified, sometimes as far back as the 4th and 5th centuries BC. It transpires that biases and fallacies are intimately related. It is standard parlance to reference a bias—understood as the psychological, or internal, and non-directly observable or ascribed part—in order to explain why a fallacy—as the behavioral, or external, and directly observable part—came about. Generally speaking, the recent empirical study of human reasoning under controlled experimental conditions has delivered important quantified results. Such may well count as historical novelties, but the phenomena are certainly much older.

Although scholarly disagreement continues to persist with respect to the interpretation of various "classical" reasoning tasks—such as the "Linda problem" or Wason's selection task (see, e.g., Charness et al., 2010; van Lambalgen, 2013)—, it is presently considered well-established that humans by and large overestimate the extent to which their beliefs are true, correct, warranted, or reliably acquired *vis-à-vis* alternative beliefs or belief-forming methods (confirmation bias). In some cases, agents display behavior consistent with the assumption that they not only ignore, but do not even perceive counterevidence; in other cases, should counter-evidence register, this may occur for the sole-purpose of formulating rebuttals, the persuasive strength of which may, in turn, be estimated incorrectly. Although the empirical evidence may still count as "somewhat sketchy,"

> [a] few studies have looked at people's perceptions of specific influences that compromise objectivity. Studies have shown people's inclination to view those whose opinions differ from their own as influenced by self-interest (Reeder, Pryor, Wohl, & Griswell, 2005), personal affections (Frantz, 2006), political partisanship (Cohen, 2003), an inability to see things fairly (Ehrlinger, Gilovich, & Ross, 2005), and unwavering ideology (Robinson et al., 1995). (Kennedy & Pronin, 2008, p. 834)

Further evidence includes experimental studies that, by and large, even see physicians report preferring a treatment with an 80% success-rate over a different treatment with a 20% failure-rate. This is interpreted as a "framing effect." After all, despite otherwise identical content, the positively framed version of a message is demonstrably preferred over the negative version. If experimental results transfer, then similar effects hold both for everyday and professional judgment and action.

3. FALSE POLARIZATION

A standard teaching and learning activity (TLA), suitable, e.g., for an introductory level course in critical thinking, might familiarize learners with a selection of biases. This normally occurs by way of presenting written or transcribed examples of externalized reasoning episodes (aka arguments) exemplifying one or more biases. A typical assignment might consist in documenting additional examples from the media, or from personal interaction. Such TLAs require applying a classification schema in order to identify bias instances, much in the way that students might apply a botanical taxonomy to identify plants in their neighbor's backyard.

A shortcoming of TLAs that are designed in a similar "show & tell"-fashion, or so Fischhoff (1982) demonstrates, is the unilateral improvement of the average ability to discern biases ascribable to (the reasoning of) a third party. These TLAs may be expected to increase the frequency of (correct) third party bias ascriptions, yet fail in improving the ability to discern own biases. In the above metaphor, students thus fail to apply their knowledge to the plants in *their own* backyard.

A social phenomenon sustained by such failure is known as *false polarization* (Keltner & Robinson, 1993, 1996; Monin & Norton, 2003; Robinson, Keltner, Ward, & Ross, 1995). The term refers to disputants' tendency to overestimate the degree of their disagreement, i.e., "how far apart the parties are" (Figure 1). Importantly, one

may grant that disputants have good reasons to disagree in the first place (see Zenker, 2012, Sect. 5). The issue, then, is *not* that groups or individuals polarize, but that their perceived difference over some issue—when expressed as a comparative distance (ordinal level)—is greater than their *de facto* difference. See Sunstein (2002) for the "true" polarization phenomenon, where a group starts out with (a joint belief in) some position, *P*, on an issue, then discusses, and ends with a (joint) belief in a more extreme position, *P**, or with a greater degree of (joint) belief in *P*.

Figure 1: Overestimating the Distance of Disagreement (from Pronin, Puccio, & Ross, 2002). The upper case illustrates an ordinal distance according to actual views. In the lower two cases, opponent groups (triangles; circles) overestimate their comparative distance.

As Kennedy and Pronin (2008) suggest, an apparently greater disagreement invites the imputation of greater bias, and *vice versa*. They also suggest that bias ascription correlates with opponents' mutual expectations that the disagreement will escalate, rather than be resolved through cooperation. Hence, "as viewpoints become more radically opposed in a discursive context, the prospects worsen for finding common ground from which either disputant may be engaged in productive discussion" (Kenyon, forthcoming). Such overestimation, then, adds a gap to the true distance between proponent and opponent position.

Puccio & Ross (1998), and Pronin, Puccio, & Ross (2002) suggest on the basis of experimental evidence that a reliable way of improving the ability to self-discern own biases—and, thus, mitigate or avoid false polarization—consists in considering the strengths and weaknesses of positions and reasoning not endorsed personally (aka "seriously considering the other side"). Also see Larrick (2008, p. 323f.)

> Partisans in the express-own-position condition in these studies showed the expected false polarization effect, markedly overestimating the gap between the positions of the two sides. By contrast, participants in the express-other-position (and, in one study, those in a third condition in which they expressed both positions) hardly overestimated this gap at all. (Pronin, Puccio, & Ross 2002, p. 653)

Standardly, cognitive and social barriers are cited to explain the false polarization effect. As with other biases, humans appear to suffer from a "bias blind-spot," in the sense of entertaining the belief that a bias is affecting others, but not themselves (ibid.). Consequently, knowing *about* a bias will not suffice to mitigate own biases. Possibly, there is a temptation to believe that introspection or will-power alone is efficient for de-biasing. However, as Frantz (2006) and Hirt and Markman (1995) submit, introspection and mental effort may, in some cases, even make a bias worse. Some authors further hold that a merely cognitive effort may, in particular cases, remain insufficient for de-biasing (Frantz & Janoff-Bulman, 2000). Perhaps an emotional component, or a quasi-religious experience, might be required.

More specific to false polarization, under normal conditions there are social barriers which may, in the presence of opponents or in-group members, keep disputants from openly admitting that they are, or have been, prone to biases. After all, granting reasonable elements of an opponent's position may appear disloyal, or hostile, to one's in-group. In doing so, one may further provide the material for an opponent's *ex concessis* attack—widely thought to be the most persuasive form of counter-argument. Finally, one may appear weak-minded to observers or audiences.

We now lay out a way for these barriers to be potentially overcome in the context of class-room instruction.

4. FOSTERING THE ABILITY TO SELF-DISCERN BIASES

From learners, this teaching and learning activity (TLA) demands engaging in 'counterfactual meta-cognition', i.e., engaging in reasoning episodes that one does not agree with personally. Required from instructors is an above average ability to (prepare teaching materials that) present a balance of reasoning and argumentation originating at both sides of an issue. Typical examples are found in social policy controversies (e.g., Zenker, 2010 on human embryonic stem cell research). A background in discourse analysis and informal logic may be found indispensable. The success of this TLA will also depend on personal qualities of the instructor, notably situational knowledge, trustworthiness, and neutrality (see Section 6). Generally, the following can be situated in the (age old) 'Forum Movement'-tradition (Keith, 2007), and may claim support from research results obtained in social and cognitive educational psychology laid out, for instance, by Kuhn (1992).

4.1 Analytical description

The TLA (i) assesses the position a learner holds on some issue (by indicating distances, as in Figure 1, above), (ii) then assigns the learner to a group of similarly opinioned students who discuss arguments supporting an opposite or, at any rate, a different position, in order to (iii) better estimate the strengths and weaknesses of these positions (individually and through teacher feedback), before (iv) learners report that position, and (v) compare it to their initial distance estimate.

On the basis of (v), the identification of biases in the initial assessment may be pursued. It then becomes a precondition that, prior to this TLA, students acquire some familiarity with the bias literature. In terms of decreasing the distance of

disagreement, however, learning does not seem to depend on such familiarity.

4.2 Sequential description

Stepwise, with structural-learning terms (McKeachie & Svinicki, 2006) bracketed (steps 1 and 8 require dedicated sessions):

1. Standard presentation of the bias literature with examples and discussion.
2. The group agrees on a topic that students potentially polarize over, such that subgroups for the most part either favor or disfavor a position (such as capital punishment, genetic modification, just war, nuclear energy, gun control, etc., or more mundane topics such as "The new university policy on …").
3. Group members call out reasons pro/con one or the other position (*pre-structural*). Reasons are noted (on the blackboard or via electronic means) under two columns: favor (pro) vs. disfavor (con) (*uni-structural*).
4. The group provides an initial structure over these reasons by distinguishing reasons that are motivated by value-considerations from those motivated by the (non-)desirability of effects (deontic vs. consequential) (*multi-structural*).
5. Members self-identify as endorsing one rather than another position, and are allocated to subgroups (of 5-7 participants, ideally). Those endorsing position X are assigned to a subgroup engaging with a position that entails: X is false. More than two subgroups can be formed, if the number of learners requires it. Members who have a neutral attitude to the issue at hand can be assigned observer-roles. They meet with a group, but do not interact argumentatively. Their function is to later provide a third party report on the group interaction. In particular, they should record which reasons were (not) discussed intensely.
6. Subgroups discuss the reasons made available earlier and relate pro-reasons with con-reasons, so as to create a structure in which a pro-reason should "meet" with a dialectically adequate con-reason (e.g., one reason may provide a rebuttal or a critique of, or a rejoinder, or a follow-up claim to, another reason). Learners can both generate and check their pairings by entertaining questions such as "Why might a reasonable person believe the opposing view?" or "If I believed the opposing view, what would my best reasons be?" Subsequently, reasons are added, and may be strengthened/weakened through additional information (*relational level*). This can occur through input from the instructor, or from other learners. Subgroups then present the position they have engaged in to subgroup members.
7. Subgroups return to the plenum, and present their order and structure over pro/con reasons, pointing out new information that was developed in the subgroup. The structures are compared and

discussed, to "pick the best of each," i.e., arrive at a joint structure, or a similarity class thereof. (Argument diagramming software may prove helpful here.) This serves to observe which reasons have *not* met with an adequate con-reason, thus identifying the strongest reasons pro/con an issue, on the assumption that a given reason—provided other qualities such as internal consistency—may not be assigned a non-arbitrary defeater, including considerations of moral value (*abstract level*).

8.   In a dedicated session, at least one day after the previous, learners return to their initial assessment of the positions endorsed, and seek to re-assess if and why they initially considered reasons to be, in a relevant sense, stronger than they turned out to be. Naming biases becomes conditionally relevant here. After all, students may wish to provide to themselves some explanation why they had initially thought of some position as stronger or weaker. Citing own biases provides such an explanation. Based on this re-assessment, students should be able to identify which of a set of biases may account for their initial assessment. Hence, it is not so important to "get the name of a bias right," but to enable students to comprehend that a bias may be at work in their own reasoning.

5. LEARNING OUTCOMES CROSSED WITH ASSESSMENT TECHNIQUES

Assessment techniques (Table 1) are chosen such that (i) instructors can provide feedback on the initial ("preparatory") and the final learning outcome; (ii) peers (subgroups) provide intermediate feedback, followed by (iii) group feedback. Discerning one's own biases is an individual task, and so is collecting evidence of biases in natural language discourse. The minute papers serve to "get the facts straight." The more important assessments involve feedback on the application and transfer of knowledge, and discussion, to discourage surface-learning.

Learning Outcome Successful Learners will be able to …	Minute Paper	Field Observation, Presentation, Feedback	Essay with Peer Review	Individual Work, Oral Presentation, Feedback	Group Work, Oral Presentation, Feedback
1. Report standard explanations why human reasoning is prone to biases.	✔			✔	
2. Describe the character of such biases in your own terms, and give examples.			✔		✔
3. Identify biases on the basis of third party argumentative language, written or spoken.		✔		✔	
4. Express at least one standpoint on the following question issue: …?	✔		✔		
5. Apply your new knowledge to your own standpoints, by identifying one on which you may (have) entertain(ed) a biased view. (Name the bias.)		✔		✔	

The top-left cell spans: Assessment Technique

Table 1: Assessment Techniques crossed with learning outcomes

6. DISCUSSION

In simple terms, the immediate purpose of the TLA is for learners to become more familiar with the strengths and weaknesses of their own positions, by becoming aware of positions they do not endorse personally. This is here called "engaging with the other side." Through such engaging, one's own biases may be better discerned than without such engaging, or so is the central assumption.

The more general goal is to provide students with two cognitive benefits: (i) an awareness of the insidiousness of bias, according to which an initial sense of being unbiased should not preclude taking steps to de-bias anyhow; and (ii) a model for de-biasing that implicates charitable engagement with interlocutors' views, rather than a mere self-check of one's presumably good intentions.

Besides character traits (e.g., intellectual courage), learning success depends on the quality of the argumentative material made accessible to learners (see below), and on how success is measured. As for measurement, it is here considered relatively unimportant that learners acquire the entrenched term for some biases, as these are used in the literature. In principle, learners may very well pick their own names, then compare those—through assistance from the instructor—to the technical term(s). Hence, success should not be measured by the student's ability to correctly repeat terms, but by the ability to identify the "character" of a bias. At the same time, should a significant number of students not be able to self-discern a single bias in their own initial assessment of a position they endorse personally, the TLA should likely count as a failure.

Secondly, the cognitive availability of plausible counterfactuals is a partial function of one's actual commitments. So, learners may not readily succeed at

contemplating reasons which support a view incompatible with their own: "[F]inding such reasons is likely to be difficult, and the judge [or agent] may infer from this difficulty that there are not many—or else they would not be so difficult to generate" (Sanna et al., 2002). Instructors will need to prepare for such difficulty, i.e., prepare to deliver reasons unknown to learners, in ways that will be found accessible.

After all, it cannot reasonably count as a learner's "fault," should instructors withhold on relevant considerations. To give but one example, to reasonably engage with the nuclear energy-issue, a long term perspective (of several generations) and the rate of radioactivity decay better be mentioned. This recommends using material that is already studied to some extent, for instance in controversy studies (van Eemeren et al., 2009)—a strand of discourse analysis—or in better quality critical thinking textbooks (e.g., Kenyon, 2008).

As indicated above, visual representations of debates ("controversy maps") may be provided. Some—e.g., scientific controversies such as the corpuscle vs. wave interpretation of light—will not be suitable in all conditions. Generally, a trade-off may be unavoidable between "keeping facts straight" and keeping learners engaged in issues they take positions on. See Newell, Beach, Smith, & Van Der Heide (2011).

7. SUMMARY

This TLA relies on empirical evidence that standard ways of teaching the heuristics and biases literature are unlikely to improve learners' abilities at discerning their own biases, *ceteris paribus*. We described a TLA geared towards such improvement. Its success depends on learners engaging cognitively with the reasoning of another party in an instructor-guided manner. The TLA can potentially help groups that disagree on some contested issue avoid the effects of false polarization.

It is less important that learners acquire knowledge of a bias' technical term, though the TLA can be adapted to this outcome. Having stressed the role of a qualified teacher, learning outcomes and assessment techniques were proposed, and sources for teaching material pointed out.

The instructor's most important tasks consist in the development/adaptation of this material, and the "fine tuning" of the TLA to students' needs and prerequisites over a total of three deployments. This includes determining the appropriate time the TLA shall take, as well as the preceding and subsequent TLAs, e.g., a lecture and a reflection-unit.

REFERENCES

Charness, G., Karni, E., & Levin, D. (2010). On the conjunction fallacy in probability judgment: New experimental evidence regarding Linda. *Games and Economic Behavior*, 68, 551–556.

Cohen, G. L. (2003). Party over policy: The dominating impact of group influence on political beliefs. *Journal of Personality and Social Psychology*, 85, 808–822.

Conley, T. M. (1990). *Rhetoric in the European Tradition*. Chicago and London: The University of Chicago Press.

Eemeren, F. H. van, & Garssen, B. (2009). *Controversy and Confrontation*. Amsterdam: John Benjamins.

Ehrlinger, J., Gilovich, T., & Ross, L. (2005). Peering into the bias blind spot: People's assessments of bias in themselves and others. *Personality and Social Psychology Bulletin*, *31*, 680–692.

Fischhoff, B. (1982). Debiasing. In D. Kahneman, P. Slovic, P., & A. Tversky (Eds), *Judgment Under Uncertainty: Heuristics and biases* (pp. 422–444). Cambridge: Cambridge University Press.

Fischhoff, B. (2010). Judgment and decision making. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(5), 724–735.

Frantz, C. M. (2006). I AM being fair: The bias blind spot as a stumbling block to seeing both sides. *Basic and Applied Social Psychology*, *28*, 157–167.

Frantz, C., & Janoff-Bulman. R., (2000). Considering both sides: The limits of perspective-taking. *Basic and Applied Social Psychology*, *22*, 31–42.

Gigerenzer, G., Todd, P.M., & the ABC Research Group (1999). *Simple heuristics that make us smart*. New York; Oxford: Oxford University Press.

Gigerenzer, G., and Brighton, H. (2009). Why biased minds make better inferences. *Topics in Cognitive Science, 1*, 107–143.

Hamblin, C. (1970). *Fallacies*. London: Methuen.

Hirt, E. R., & Markman. K.D. (1995). Multiple explanation: A consider-an-alternative strategy for debiasing judgments. *Journal of Personality and Social Psychology*, 69, 1069–1086.

Kahneman, D. (2011). *Thinking fast and slow*. New York: Farrar, Strauss and Giroux.

Kahneman, D., Slovic, P., & Tversky, A. (Eds) (1982). *Judgment Under Uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.

Keltner, D., & Robinson, R.J. (1993). Imagined ideological differences in conflict escalation and resolution. *The International Journal of Conflict Management*, *4*, 249–262.

Keltner, D., & Robinson, R.J. (1996). Extremism, power, and the imagined basis of social conflict. *Current Directions in Psychological Science*, 5, 101–105.

Kennedy, K. A., & Pronin, E. (2008). When disagreement gets ugly: Perceptions of bias and the escalation of conflict. *Personality and Social Psychology Bulletin*, *34*, 833–848.

Kenyon, T. (2008). *Clear Thinking in a Blurry World*. Scarborough, ON: Nelson.

Kenyon, T. (2013). False polarization as applied social epistemology. *Synthese* (forthcoming).

Keith, W. (2007). *Democracy as Discussion: Civic Education and the American Forum Movement*. Lanham, MD: Lexington Books, 2007

Kuhn, D. (1992). Thinking as argument. *Harvard Educational Review*, *62*(2), 155–178-

Lambalgen, M., van (2013). The descriptive complexity of the Wason selection task is Σ^1_3 and other tall stories. In M. Urbanski (Ed.) *Title to be announced* (Proceedings of Logic and Cognition, Poznan, Poland, May 2012). FoLLI Publications on Logic, Language and Information. Dordrecht: Springer (forthcoming).

Larrick, R. P. (2008) Debiasing. In D. J. Koehler, & N. Harvey (Eds), *Blackwell Handbook of Judgment and Decision Making* (pp. 316–337). Malden, MA: Blackwell Publishing Ltd.

Monin, B., & Norton, M. (2003). Perceptions of a fluid consensus: Uniqueness bias, false consensus, false polarization, and pluralistic ignorance in a water conservation crisis. *Personal and Social Psychology Bulletin*, *29*(5), 559–567.

Newell, G. E., Beach, R., Smith, J., & VanDerHeide, J. (2011). Teaching and Learning Argumentative Reading and Writing: A Review of Research. *Reading Research Quarterly*, *46*(3), 273–304.

McKeachie, W. J, & Svinicki, M. (2006). *Teaching Tips. Strategies, Research and Theory for College and University Teachers*. Houghton Mifflin Company: Boston & New York.

Pronin, E., Puccio, C., & Ross, L. (2002). Understanding misunderstanding: Social psychological perspectives. In T. Gilovich, D. Griffin, & D. Kahneman (eds), *Heuristic and Biases: The Psychology of Intuitive Judgement* (pp. 636–665). Cambridge: Cambridge University Press.

Puccio, C., & L. Ross (1998). Real versus perceived ideological differences: Can we close the gap? *Unpublished Manuscript*, Stanford University.

Reeder, G. D., Pryor, J. B., Wohl, M. J. A., & Griswell, M. L. (2005). On attributing negative motives to others who disagree with our opinions. *Personality and Social Psychology Bulletin*, *31*, 1498–1510.

Robinson, R., Keltner, D., Ward, A., & Ross, L. (1995). Actual versus assumed differences in construal: "Naive realism" in intergroup perception and conflict. *Journal of Personality and Social Psychology*, *68*, 404–417.

Sahlin, N.-E., Wallin, A., & Persson, J. (2010). Decision Science: from Ramsey to dual process theories. *Synthese*, *172*, 129–143.

Sanna, L., Stocker, S., & Schwarz, N. (2002). When debiasing backfires: Accessible content and accessibility experiences in debiasing hindsight. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(3), 497–502.

Sunstein, C. R. (2002). The Law of Group Polarization. *Journal of Political Philosophy, 10*(2), 175–195.

Tversky, A., & Kahneman, D (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*, 1124–1131.

Zenker, F. (2010). Analyzing Social Policy Argumentation: A Case Study of the 2007 Majority Opinion of the German National Ethics Council Regarding an Amendment of the Stem Cell Law. *Informal Logic*, *30*, 62–91.

Zenker, F. (2012). The Explanatory Value of Cognitive Asymmetries in Policy Controversies. In J. Goodwin, & R. Delaplante, R. (Eds), *Between Scientists and Citizens* (Proceeding of the First Int. Conference of the Great Plains Society for the Study of Argumentation, GPSSA, Ames, Iowa, USA, June 1-2, 2012) (pp. 441–452). Ames, IO: GPSSA.