The role of mental states in argumentation: Two problems for rationality from the psychology of belief

DAVID GODDEN
Department of Philosophy
Old Dominion University
Norfolk, Virginia
U.S.A. 23529
Email: dgodden@odu.edu
www.davidgodden.ca


ABSTRACT: This chapter recognizes the contributions made to the theory of argument by the work of Cristiano Castelfranchi, together with Fabio Paglieri, by situating their work in the development of social, or process-based accounts of argumentation. It is argued that this orientation to the social requires grounding in the psychological, and thus calls for a belief-based perspective on argumentation. It is shown how Castelfranchi’s work on the ontology of belief in relation to goals and intentions, together with the Data-oriented Belief Revision model contributes to this approach by bridging the gap between the social and the psychological. The paper concludes by raising two problems for standard models of argument arising from the psychology of belief: (i) that we seem to lack adequate voluntary control over our beliefs to be rationally responsible for them, and (ii) that we seem not to be reason trackers in the way required by standard accounts of rationality employed in argumentation.

KEYWORDS: argumentation; belief perseverance; belief revision; Castelfranchi, Cristiano; cognition; Data-oriented belief revision; deontological theory of rationality; psychology of belief; rationality; voluntarism argument
1. INTRODUCTION

In pioneering work, Castelfranchi and, more recently, Castelfranchi and Paglieri, have brought together two parallel but previously disparate research traditions: formal theories of rational belief revision and theories of argumentation understood as inter-individual systems of commitment management and rational dispute resolution. In doing so they have brought argumentation theorists to recognize the central role of belief, and other doxastic attitudes, in argumentation. By and large, formal theories of belief revision consider beliefs as theoretical entities – components of constructed formal systems. Yet, beliefs are mental, psychological states of boundedly-rational agents, and this presents a variety of problems for belief-oriented theories of argumentation.

After making the case for the need for belief-centered theories of argument, the paper considers the contributions that the formal theory of belief revision, specifically Castelfranchi and Paglieri’s Data-oriented Belief Revision (DBR) theory, has made to argumentation. The paper then considers two, related problems arising from the psychology of belief, which have important consequences for theories of rationality and belief revision, as well as any theories of argumentation connected thereto: (i) doxastic voluntarism, the causal efficacy of reasons, and our rational accountability for belief; and (ii) our cognitive capacities as reason trackers.

A brief comment on each is in order. On the deontological conception, to be rational is to be epistemically responsible. A rational agent is one who can be held accountable for her doxastic attitudes. She is rational to the extent that her doxastic attitudes are rightly connected to reason. Yet, our beliefs are not the sorts of things over which we have conscious, voluntary control. Thus, it is difficult to understand how believers can be held rationally responsible for them.

Secondly, to rightly connect our beliefs to reasons, it would seem that any target doxastic attitude ought to change correlativelly with, and proportionately to, changes in the relevant evidential beliefs. Yet beliefs are not causally connected to their reasons; the evidential relations that occur amongst our beliefs are not mirrored at a causal, psychological level. So, it seems that, on the deontological conception of rationality, we must be reason-trackers. Indeed, Brandom (1994, 2000) contends that in order to be sapient, concept-users we must be scorekeepers (i.e., commitment and entitlement trackers) in the game of giving and asking for reasons. Yet, empirical research indicates that our capacity for reason tracking does not bring us close to the standards of epistemic rationality to which we typically hold ourselves accountable in normal argumentative practice.

I contend that these psychological features of belief, and issues consequent to them, present significant obstacles for theories of argumentation and belief revision that aspire to provide a normative component.
2. AN HISTORIC PERSPECTIVE SHIFT IN THE STUDY OF ARGUMENT

Argumentation studies, in its present state, may be viewed as a disciplinary collective which seeks to study argumentation as an essentially situated activity engaged in by autonomous, cognitive and rational agents. This disciplinary collective emerged from a set of reactions to a logical approach to argument, which conceived of its subject matter as a purely semantic entity. Arguments were defined as sets of propositions, one of which was designated as the conclusion the remainder of which were the premises. These abstract entities were interpreted and evaluated purely on the basis of their formal and semantic properties of premise truth and inferential validity. Pragmatic and situational properties were considered descriptively and normatively irrelevant and deliberately ignored. For example, arguments were interpreted and evaluated independently of their (i) intentionality (purpose or function) – e.g., that the argument was offered in an effort to rationally persuade, (ii) agent-relatedness (or audience) – e.g., that the propositions are instantiated by arguers as speech-acts directed to audiences, and (iii) situation (or context) – e.g., that the argument was offered as a in a public, institutionalized setting rather than a private, personal one. Such an approach not only neglected the rhetorical dimensions of argument, but its pragmatic dimensions as well.

In the late 1970’s, a fundamental perspective shift occurred in the study of argument. This shift may be roughly identified by Brockriede’s (1975, p. 179) methodological hallmark that “people will find arguments in the vicinity of people.” Responding to O’Keefe (1977), Brockriede (1977) distinguished between two senses of “argument,” as (i) product (argument1) and (ii) process (argument2) to which Joseph Wenzel added a third perspective of procedure (1979). Rather than being conceived of as abstract entities essentially devoid of contextual features, arguments were now conceived of as essentially situated artifacts deployed by agents engaged in purposeful activities.

Not only did this occasion a methodological shift in focus from product-oriented, to process-oriented approaches. Additionally, the very subject matter of argument was recast. Arguments were no longer abstract semantic objects (sets of propositions), but doings (or attempts at doings); they were communicative ac-

---

1 O’Keefe (1977) distinguished two concepts of argument:
(i) argument1, or “arguing-that” which is “a kind of utterance or a sort of communicative act;” and “something one person makes (or gives, or presents or utters).”
(ii) argument2, or “arguing-about” which is “a particular kind of interaction;” “something two or more persons have (or engage in).”

2 Wenzel(1979, pp. 115-116) paired these three perspectives to the Aristotelian branches of the study of argument as follows: logic (product-based), rhetoric (process-based), and dialectic (procedure-based).
tions. Arguments, on this new conception, became activities defined according to their goals (typically demonstrative or persuasive) and their means (typically rational).

In addition to characterizing arguments pragmatically rather than semantically, the perspective shift fundamentally humanized the study of argument by focusing on the inherently social or interpersonal aspect of argument. From a rhetorical perspective, Willard (1989, p. 1), claimed “Argument is a form of interaction in which two or more people maintain what they construe to be incompatible positions.” Dialectical perspectives defined argument in terms of a verbal, social activity. For example Walton (1990, p. 411) defined an argument as “a social and verbal means of trying to resolve, or at least contend with, a conflict or difference that has arisen between two parties engaged in a dialogue.” Similarly, from the Pragma-Dialectical perspective: “Argumentation is a verbal and social activity of reason aimed at increasing (or decreasing) the acceptability of a controversial standpoint for the listener or the reader, by putting forward a constellation of propositions intended to justify (or refute) the standpoint before a rational judge” (van Eemeren et al. 1996, p. 5). Even traditionally product-based approaches, such informal logic, embraced a dialectical approach which conceived of the argument-product as the outcome of a situated practice. For example, Blair and Johnson (1987, p. 45) wrote that “An argument understood as product … cannot be properly understood except against the background of the process which produced it – the process of argumentation” (cf. Johnson 2000, p. 12).

Clearly this shift from product to process affected the interpretative task of argument analysis, which now focused on pragmatic doings (speech acts) rather than semantic things (sentences). Perhaps more importantly, concomitant changes also occurred in the normative task of argument evaluation. Process-based, dialectical approaches to argumentation reconceived argument norms as essentially proce-
dural rather than structural (e.g., logical or epistemic). Johnson and Blair (1994, p. 13) describe this shift as follows:

The dialogue logician assigns to logic the task of prescribing rights and duties in the transaction of a rational dialogue. The [product-based] informal logician assigns to logic the task of developing the criteria or standards for use in the evaluation of arguments. (cf. Johnson, 2000, p. 291)

Typically these procedural norms were taken to derive (a) from the type of activity (read “argumentative discussion”) engaged in – specifically the overall goal or purpose ascribed to that activity – and (b) from the type of act performed in the course of that activity (Hamblin 1970; Walton & Krabbe 1995; Walton 1998). For example in Pragma-Dialectics (van Eemeren and Grootendorst 2004, pp. 17, 132) the rules governing a critical discussion (a normative ideal for argumentative dialogues which seek to rationally resolve a difference of opinion) are justified by their problem validity (an instrumental standard of how successfully the rules bring about the goal of the discussion) and intersubjective validity (an anthropological standard of reasonableness, external to any particular instance of argumentation). Fallacies, on such accounts, are conceived of as discussion moves which impede the resolution of the discussion.

By contrast, the design approach (Goodwin 2002; 2007) to argument does not presuppose or impose an overarching purpose upon the activity (read “dialogue”) itself, but instead focuses on the goals of individual, situated and autonomous actors acting strategically. Instead of as a rule-governed dialogue, design theory (Goodwin 2002, p. 6) conceives of the activity of arguing as (paradigmatically) a transaction – understood (as per Webster’s) as “a communication or activity involving two parties or two things reciprocally affecting or influencing each other.”

Argumentative norms, on this view, are located and applied from a bottom-up, rather than a top-down, approach.

Instead of deriving norms from the standards set by some social function external to arguing, design theories contend that the argumentative transaction is internally self-regulating. Each arguer, to achieve her goals, tries to establish for herself and the other participants a normative environment within which their arguing can proceed. (Goodwin 2002, p. 6)

On the design approach, norms come to bear on argumentative situations as arguers act so as to manifestly undertake, discharge, impose and enforce certain commitments (or responsibilities) attached to the ordinary pragmatics of speech acts.

3. THE LEGACY OF COMMITMENT IN ARGUMENTATION

One of the legacies of the shift to the pragmatic, dialectical, process-based perspective on argumentation, where the basic components of arguments are viewed
not as a proposition, or sentences (e.g., premises or conclusions), but speech acts, or utterances (e.g., moves in a language game or dialogue), has been that commitment has become the central locus of argument interpretation and evaluation. Commitments arise out of, and are attached to, the public doings of arguers, and have become not only the units in which argumentative positions are described (since utterances both instantiate and generate commitments), but the markers or argumentative responsibilities.

Commitments are understood as “proposition[s] one has gone on record as accepting [i.e., verbally (see below)]” (Walton 2010, p. 23). So, in one sense, commitments are simply speech acts – bits of verbal behavior. Yet, they also have an inherently normative dimension which is cashed out as sets of responsibilities one undertakes in respect to some particular claim. Commitments, in this sense, are obligations. Agents incur these obligations by doing certain kinds of things, specifically by making speech acts in the right sorts of circumstances. For example, if an agent asserts that c, she has undertaken a set of responsibilities to do things like (i) support c with sufficient and acceptable reasons if challenged, and (ii) retract c in the event that she cannot meet her first responsibility. Similarly, the commitments that arise from conceding that c might include things like the following: (i) revising one’s other commitments such that they are consistent with c, (ii) not making assertions one takes to be inconsistent with c, (iii) allowing c to be used as a premise, and (iv) not arbitrarily abandoning c.

Importantly, these responsibilities are inherently public and social. They derive not only from the nature of the speech act itself, but often also from the particular rules governing the activity (read “argumentative discussion”) in which one is engaged at the time of making the speech act. In addition to the overarching telos of the argumentative discussion (if any), these commitments also serve – at the local level – as the locus, and sometimes even the basis, for argumentative norms, understood as procedural obligations arguers have to themselves and one another.

As Walton (1998, p. 31) writes, “The concept of commitment is the basic idea behind all dialogue as a form of reasoned argumentation.” This focus on commitment can be found in formal approaches to argument dialogues (e.g., Hamblin 1970), pluralistic goal-driven and speech-act approaches (Walton & Krabbe 1995, Walton 1998), highly idealized and normalized approaches to argumentative dialogue (such as the Pragma-Dialectical approach; van Eemeren & Grootendorst 1984, 2004), and even design approaches (Goodwin 2002, 2005). By focusing on the essentially social and behavioral aspects of argumentation, the perspective shift to the process-based view of argument has resulted in a commitment-oriented view of it.

4 In one sense, the utterance-to-commitment relationship is many:one, since there can be many different tokens for , or even ways of instantiating, a particular commitment. In another sense, there is a one:many relationship between utterances and commitments, since a single utterance can generate several different commitments (see below).
While humanistic, this process-based, commitment-oriented conception ignores the cognitive and psychological dimensions of argument. It is not only that the producers and consumers of arguments are psychological as well as situated and social beings. More specifically, their doxastic life has both a causal and an explanatory relationship to their social and behavioral life, and as such it seems both primary and ineliminable to a properly humanized study or argument. Persuasive argument seeks to affect its audience’s behavior by changing their minds with reasons. As such, arguments contain or represent inferences, and the stakes in argument are cognitive attitudes. Thus it would seem that the move to the social ought to have resulted in a move to the psychological. These ideas is explored below in the context of the role belief should play in explaining and evaluating argumentative behavior, and in analyzing and appraising arguments.  

4. THE PRIMACY OF BELIEF IN ARGUMENTATION

4.1 Belief and Acceptance

Godden (2010, p. 399) proposed an account of belief with minimal metaphysical ‘freight,’ on which beliefs have two defining characteristics. First, beliefs have what Searle (1979, pp. 7-9) called a world to mind fit; second, they are what Ramsey (1931, p. 28) called the maps by which we steer. Together, these properties are sufficient to distinguish beliefs from commitments and acceptances.

Acceptance comes in two varieties: verbal acceptance (conceding), understood as a speech act of assent, and mental acceptance (accepting proper). Pinto (2003, p. 8) describes (mentally) accepting a proposition as “being prepared to use it as a premiss in my reasonings or inferences (or in cases of public discussion being prepared to tolerate its use by others).” This account roughly coincides with Cohen’s (1992, p. 4) definition: “to accept that \( p \) is to have or adopt a policy for deeming, position, or postulating that \( p \)” (cf. Stalnaker 1984, pp. 79-80).

Paglieri and Castelfranchi (2006a) distinguish belief and acceptance according to the “different functional roles … [they] play in the cognitive economy of the subject.” Belief has an alethic function which “is meant to provide a veridical representation of the world,” while acceptance has a pragmatic function, “its role is to provide a representation of the world that is suitable for supporting successful deliberation and effective action.” While belief and acceptance are not exclusionary (indeed in normal cases they might coincide), they are not coextensive either.

For Cohen (1992, p. 4), beliefs are dispositions of a certain sort; they inclinations towards feeling that a certain proposition is true: to believe that \( p \), is to, upon introspection, “normally … feel it true that \( p \) and false that not-\( p \).” As such, our beliefs play a premissory role in our inferences and practical reasoning, they

---

5 Belief is here intended only to be a representative of the conscious mental attitudes which are advanced as being the proper focus of theories of argument (see Godden 2010, p. 398, fn1).
have a causal role in determining our actions, and they contribute significantly to a rational explanation of our behavior. Davidson (1980, pp. 4-5, fn. 2) describes actions as the intentional doings of agents – acts which, under some description, can be said to be done for a reason. Roughly, then, to attribute actions to agents is to take what Dennett has called the intentional stance towards them; this involves explaining their behavior in terms of beliefs and desires which are taken to causally determine the outward behavior.

4.2 Belief as the Locus of Persuasion

Given these conceptions of belief, acceptance and commitment, Godden (2010, pp. 404-406) then argued that belief, not commitment, is the primary target at which persuasive argumentation should be directed. The argument presented there ran roughly as follows. Commitment-based theories of argument (e.g., Hamblin 1970; Walton and Krabbe 1995; van Eemeren and Grootendorst 1984, 2004) tend to share a commitment to the following three claims:

(I) Goal: The goal of persuasive argumentation is to settle a difference of opinion by rational means.

(II) Independence of belief and commitment: Commitment and belief are logically and causally independent; a change in one does not always result in a corresponding change in the other.

(III) Resolution: A difference of opinion is resolved when the commitments of the disputants have reached a state of agreement with respect to the claim at issue. (Godden 2010, p. 404)

Yet, a paradigm of failure of persuasive argumentation occurs whenever an arguer concedes a position in argument (e.g., as a speech act or as a move in an argumentative dialogue) but subsequently (perhaps following the argumentative exchange) acts as though no such concession had been made. Because such argumentation fails to determine, or even to affect, the future actions of arguers, a genuine resolution to a difference of opinion is not achieved, despite the appearance of a resolution having been achieved through the concession.

Whatever the agent’s motives for making such a concession, that it fails to subsequently inform her behavior indicates that her beliefs are not aligned with her commitments on the matter at issue. Since, by definition, an agent’s actions are shaped by her beliefs, to be effective resolutions of differences of opinion must occur at the level of belief and not merely at the level of verbal commitment. Even though we may be able, sometimes, to hold agents answerable to commitments they made in argumentation, when the agent’s own beliefs do not coincide with those commitments she well not hold herself accountable to them. Yet, that she do so is normally necessary for the resolution of a difference of opinion to be effective, since interlocutors in argumentation will not normally be present to hold each
other accountable to commitments undertaken in argumentation much beyond the argumentative exchange.

4.3 Belief as the Foundation for Action

A second argument, touched upon in that paper (p. 401, fn. 6), might now be offered in greater detail. This argument also hinges on the point that beliefs are causally related to actions. Importantly, the normal competitors to belief as the locus of argumentation and persuasion – acceptance and commitment – are each actions in Davidson’s sense: they are intentional acts, consciously and voluntarily performed, which may be described as having been done for a reason. As intentional acts, acceptance and commitment are subject to intentional explanation, and such explanations are given in terms of the agent’s beliefs and desires.

Insofar as acceptance is like a Cartesian judgement – the conscious and deliberate endorsement of a claim – an agent’s acceptance of a claim is both explained by and justified in terms of her beliefs about the acceptability (be it alethic or pragmatic) of that claim. Similarly with speech acts, such as assertion and concession. That an agent asserts, concedes or retracts a claim in her commitment-store is explained by recourse to her beliefs concerning the relevant properties of that claim. For example, standard analytical explanations (Grice 1989; Searle 1969) of acts of assertion hold as constitutive that assertions are expressions of beliefs through uttering sentences. Thus Hindriks (2004, p. 136) gives the following condition as a criterion for assertion: “An utterance of u counts as an assertion of p just if the person who utters it expresses her belief that p.”

5. FROM THE SOCIAL TO THE COGNITIVE

The perspective shift from product to process, and from argument to arguer, reflected an attempt to humanize the study of argument and an awareness of the essential place of social and situational factors in the nature of the very subject matter under investigation in argumentation studies. Yet, it would seem that this perspective has neglected the cognitive, psychological dimensions of argument – dimensions which seem to be both ineliminable and basic.

In lamenting the lack of integration of theories of argumentation with theories of belief change, Paglieri and Castelfranchi (2005, pp. 359-360) portray belief revision and argument as distinguished only according to audience, with both aiming at what Harman (1986) has called reasoned change in view. One useful way to conceive of the activity of arguing which includes its essential, cognitive dimension is found in Campolo’s (2005, p. 41) idea that arguing is reasoning together. This conception emphasizes not only the interpersonal and cognitive aspects of argumentation, but views argument as an activity whose function is inherently reflective, reparative or remedial. The activity of reasoning is invoked when some other activity in which we are otherwise smoothly engaged is somehow interrupted. Like any other activity, our reasoning abilities (and therefore our argumenta-
tive abilities) are founded in training (socialization) as well as capacity (natural endowments). Each of these factors serves to further set limits to the prospects of reason. “[R]easoning together,” Campolo reminds us (p. 41), “is not some sort of magically creative act that always produces efficacious results. It is rather a way of drawing on shared resources, and as those resources get thinner, reasoning loses traction.” This conception begins to raise important questions concerning the effect that the cognitive limitations of arguers might have on not only the efficaciousness of argument, but also on how argumentative norms ought to be formulated and relate to arguer’s cognitive capacities.

6. FORMAL THEORIES OF BELIEF REVISION

The increased appreciation of the cognitive bases and ineliminable doxastic dimensions of the social and behavioral aspects of argumentation has provided one catalyst for the re-introduction of beliefs into theories of argumentation. Accepting the centrality of belief to argumentation, it is worthwhile to consider where theorists might turn to gain a working understanding of how belief operates in argumentation. One such direction is to formal models of belief revision.

In this context, the work of Castelfranchi, and more recently Castelfranchi and Paglieri [hereafter “C&P”], have contributed substantial advances in the theoretical understanding and representation of the operation of belief in argumentation. The theoretical framework adopted by C&P takes seriously the relevance of belief to argumentation argued for above. Indeed, on many occasions they have been outspoken advocates for the integration of theories of argumentation with theories of belief revision. For example, they (P&C 2005, p. 360) comment on the necessity of belief-talk in explaining the behavior—including their verbal behavior—of arguers.

Argumentation theories remain incomplete, if they cannot be grounded in belief revision models: they describe interesting dialogical patterns and their effects, but cannot explain why and how such effects are produced. Without an underlying model of belief dynamics, argumentation theories are forced to remain ‘out of the black box’.

Indeed, previously Castelfranchi (1996, p. 234) had argued that the causal connection between belief and action necessitates the incorporation of theories of rational belief (revision) into rational action theory. “A Cognitive agent, as a matter of fact, is an agent who founds his decisions, intentions and actions on his beliefs. … rationality in believing contributes to rationality in behaving. … Rational beliefs are a necessary condition for rational behaviour, since irrational beliefs are a sufficient condition for irrational behaviour.” Thus, Castelfranchi (1996, p. 235; cf. 1995) proposed the Autonomous Cognitive Agent Postulate:
It is impossible to directly modify the goals (and then the intentions and actions) of an Autonomous Cognitive Agent. In order to influence him (i.e. to modify his goals), another agent should modify his beliefs supporting those goals.

According to this postulate, the intentional, goal-directed behavior of agents cannot be properly explained or affected (e.g., through argument) without reference to, or effecting, corresponding changes or states in the agent’s belief structure.

6.1 Data-Oriented Belief Revision (DBR)

In this context, Paglieri and Castelfranchi developed a Data-oriented Belief Revision (DBR) model (Paglieri 2004), which is offered as an alternative to the AGM approach (Gärdenfors 1988). DBR effectively models a two-system, two-stage process of belief revision. The model (P&C 2005, p.80) distinguishes between data and beliefs where data are conceived of as “pieces of information gathered and stored” by a cognitive agent, while beliefs are data that “the agent considers reliable bases for action, decision, and specific reasoning tasks e.g., prediction and explanation.”

This distinction leads P&C to conceive of belief change as a two-step process. First, information received as data may have initial and direct effects on other pieces of data stored by the agent. Second, data that is subsequently endorsed as a belief (a process characterized as belief selection) may have secondary and indirect effects on other beliefs (pp. 80-81). Importantly, data-management is a process independent of belief revision, and involves processes such as (i) information update, (ii) data properties and assessment, and (iii) belief selection.

P&C (pp. 81 ff.) propose that data are selected or rejected as beliefs on the basis of four properties conceived as cognitive reasons to believe. These are: relevance, credibility, importance and likeability. Relevance and likeability measure a relation between data and goals, while credibility and importance are structural relations between data. P&C propose that the properties of credibility, importance and likeability determine the outcomes of belief selection (adoption) in algorithmic ways according to whether their combined value (which gives the strength of the datum) meets or exceeds a certain threshold, understood as a minimal threshold of belief. Information update, as a kind of data management, occurs independently of belief revision.

Belief revision can occur in two principal ways. First, it can be triggered externally as a result of information update and belief selection. Secondly, it can occur internally as a result of inference from existing (or selected) beliefs. Importantly, data are managed in ways that are categorically different than the ways in which beliefs are revised. Data are organized according to three relations of support, contrast and union (p. 84). Data are managed in a connectionist, parallel, and coherentist way, while beliefs are revised in a computationalist, serial and foundationalist manner according to rule-governed algorithms.
6.2 Applications of DBR

The DBR model was developed in the context of existing work by Castelfranchi which had already established a working notion of belief, and explored its relationship not only to intentional behavior but also to other cognitive entities informing behavior such as goals, intentions and commitments. For example, Castelfranchi (1996) contributed to an understanding of the relationship between theoretical and practical reasoning by demonstrating that there is a support relation between beliefs and goals whereby beliefs provide the reason for goals such that goals should change when supporting beliefs do, and supporting beliefs should be maintained when their goals are. This work is developed in C&P (2007) into a model of belief-based goal processing which provides the core for a constructive theory of intentions. In more recent work P&C (2010) this theme is revisited through a consideration of the strategic and instrumental, goal directed reasoning involved in deciding to argue based on a cost-benefit analysis.

Within the overall framework of the DBR model, the accomplishments of Castelfranchi and Paglieri have been remarkable, and their contributions to theories of argument impressive. P&C (2006b) show how the essentially coherentist AGM model cannot express or represent the basing relations at the core of argument – and reasoning, or inference – where conclusions based on reasons. They proceed to show how DBR can model argumentative relations when understood on Toulmin’s Data-Warrant-Claim structure. Importantly, this includes showing how DBR can accommodate defeasible reasoning, including premise defeaters, undercutting (warrant) defeaters, and direct (conclusion) rebutters.

Overall, Castelfranchi and Paglieri’s Data-oriented Belief Revision model provides a comprehensive and powerful model capable of representing not only argument as a kind of other-directed, inferential-based belief revision, but also the connection between belief and argumentatively relevant behaviors from speech acts and strategic maneuvering within and prior to argument, to an arguer’s actions following argumentation. Yet, in understanding belief as a component of a formal model, it is important to see how well the assumptions informing these models, and our theories of argument more generally, correspond to the actual operation of belief. The final sections of this paper turn to three psychological characteristics of belief and its operation which seems problematic to standard accounts of the role of belief in argument.

7. BELIEF, RATIONALITY AND REACTIVE ATTITUDES

Rationality is the basic evaluative ideal of argument, and virtually all theories of argumentation, whether rhetorical, dialectical or epistemic, suppose, either implicitly or explicitly, some theory of rationality. As Willard (1989, p. 152) observed “Not all rationality theories include argument in their definitions, but virtually all argument theories include rationality in theirs.”
Basically, to be rational is to respond rightly to reasons. That is, “rational beliefs must be based on reasons” (Brown 1988, p. 38; cf. Siegel 2004, p. 598). “To be a critical thinker,” Siegel writes, “is to be appropriately moved by reasons. To be a rational person is to believe and act on the basis of reasons” (1988, p. 32). Similarly, rational persuasion occurs when the Other is persuaded to accept a conclusion “on the basis of the reasons and considerations cited [in the argument], and those alone” (Johnson 2000, p. 150); thus persuasive success depends not only on the overall rationality of the result but on the apparentness of that rationality (ibid. and passim). What is it for our views to be based on reasons? Roughly, this occurs when our degree of commitment or attachment to our views somehow accords with the reasons we have for those views. Siegel (1997, p. 2) puts it this way: “to say that one is appropriately moved by reasons is to say that one believes, judges, and acts in accordance with the probative force with which one’s reasons support one’s beliefs, judgments and actions.” Similarly, Pinto (2006, p. 287) explains rationality along the lines of a qualitative evidence proportionalism whereby “rationality is a matter of making our attitudes towards propositions or propositional contents appropriate to the evidence which shapes them.”

Roughly, this picture amounts to a deontological conception of rationality. It is based on the idea that rational agents have a basic epistemic duty to know (relevant) truths and avoid error. This, combined with the idea that being rational — i.e., rightly connecting our cognitive attitudes to reason and evidence — is essential and instrumental in fulfilling our epistemic duties, forms the basis of our argumentative norms. As believers, we have a duty to be rational, and argumentative norms and obligations flow from these rational norms and obligations. Importantly, then, on this account, our beliefs and other cognitive attitudes are rightly subject to what Strawson (1962; cf. Jäger 2004, p. 1) called reactive attitudes; we rightly praise and blame people not only for what they do, but for what they believe.

On the picture we have been advancing, belief plays an essential role in argument. Beliefs feature centrally in explaining an agent’s behavior, including her argumentative behavior. Ultimately, beliefs are what is at stake in argumentation; arguers seek to affect their interlocutors behavior by changing their minds with reasons. Further, beliefs are rightly subject to rational evaluation and to our reactive attitudes. In some cases, irrationality or error is blameworthy and irrational believers can rightly be held accountable.

Yet, when considered as purely abstract, formal entities, as beliefs are in formal theories of belief revision, various contingent qualities of beliefs arising from their psychological nature, and the nature of believers as psychological beings, are glossed over or neglected. Regrettably, several of these qualities raise important problems for normative theories of belief revision, and hence to belief-centric theories of argumentation.
8. VOLITION, BELIEF AND RATIONAL ACCOUNTABILITY

A first problem is this. Beliefs do not seem to be subject to voluntary control. That is, as a contingent matter of fact, by and large, believers lack the ability to voluntarily form or revise their beliefs (cf. Feldman 2000, p. 670). To demonstrate: consider whether, at this moment, you can come to believe that the United States is still a colony of Great Britain. (If it will help, suppose you are offered the additional incentive of a large cash reward if you are successful) (cf. Alston (1989b, p. 122). As it turns out, this is just something that, as believers, we are unable to do. As Alston (1986, p. 196) writes:

For the most part my beliefs are formed willy-nilly. When I see a truck coming down the street, I am hardly at liberty to either believe that a truck is coming down the street or refrain from that belief … it is clear that for the most part we lack such powers.

It is not merely that we cannot form or change our beliefs purely at will, but rather that, choice seems to play no part in the process whatsoever. For the most part, belief formation and revision does not seem as though it could go otherwise than as it does.

Yet, as Alston proceeds, if believers lack basic voluntary control over their beliefs, it is prima facie implausible that they should be held accountable for them. Again Alston (1986, p. 196) writes:

All this talk [of deontological conceptions of rationality] has application only if one has direct voluntary control over whether one believes that \( p \) at any given moment. If I lack such control, I cannot believe or refrain from believing that \( p \) at will, then it is futile to discuss whether I am permitted to believe that \( p \) at \( t \) or whether I would be irresponsible in choosing to believe that \( p \) that \( t \). And it seems that we just don’t have any such control, at least not in general.

In general, this line of thinking has been named the voluntarism argument, a standard formulation of which is:

\[ \begin{align*}
P_1. \text{Doxastic voluntarism is false.} \\
P_2. \text{If doxastic voluntarism is false, then the deontological conception of rationality (epistemic justification) is false.} \\
C. \text{Therefore, the deontological conception of rationality and epistemic justification is false.} \text{ (cf. Feldman 1988, p. 237; Jäger 2004; Kim 1994, p. 282)}
\end{align*} \]

If correct, this line of argument affects not only deontological theories of rational belief (revision), but also all theories of argument which depend on such a theory. Yet, as argued above, it would seem that theories of argument do and should rely on such deontological theories.
As Feldman observes (1988, p. 238) there are two primary strategies of refutation available. One is to reject P1. Moves of this sort begin by considering the kinds of voluntary control agents can have over their actions, a standard list (cf. Alston 1989b, pp. 122 ff.; Feldman 2000, pp. 670-671) of which goes as follows in order of strength:

(i) *Basic voluntary control*: things we can just do (e.g., raise own hand)
(ii) *Non-basic immediate voluntary control*: things we can straightforwardly do by doing other things which, ultimately, will be under basic voluntary control. (e.g., opening doors, turning on lights)
(iii) *Long-range voluntary control*: things we can accomplish over time by doing other things (e.g., opening doors, turning on lights)
(iv) *Indirect voluntary influence*: things which we can affect over time by doing other things (e.g., controlling our weight)

Both sides of the debate tend to agree that we lack basic or direct voluntary control over our beliefs, and that the kind of control we have over our beliefs is indirect. Advocates of the voluntarism argument (Alston 1989b) have claimed that at best we generally have only indirect voluntary influence over our beliefs, and that this is not sufficient to claim that beliefs are appropriately subject to reactive attitudes. Kim (1994) accepts Alston’s claim concerning the degree of control we have, but claims that all that is required in order for us to be epistemically responsible is that we critically reflect on our beliefs, and this is surely something over which we have sufficient voluntary control. Although Feldman (2000, pp. 671 ff.) has argued that many ordinary beliefs are subject to non-basic immediate voluntary control (e.g., I can come to believe that the lights are on by turning on the light-switch), Jäger (2004, p. 5) has argued that this is not to the point: “The (alleged) problem is not that under different epistemic conditions, in different epistemic worlds, we could not believe otherwise than we actually do. The problem is that, in light of the grounds we actually have for and against a given belief, we cannot but adopt or reject it.” Instead, Jäger (2004) opts for a compatibilist view which claims that even though there may not be any alternate possibilities for believing in some given epistemic circumstance, it remains the case that our beliefs can be said to be voluntary so long as they can be caused in the right sorts of ways, such as on the basis of the evidence.⁶

A second refutation strategy is to reject P2, by claiming that obligations can still attach to behavior over which we have no control or alternatives. Unlike the compatibilist approach, this move rejects the *ought implies can* principle, and claims that we can be held responsible even in cases where we cannot meet our obligations (regardless of whether we could have done other than what we did). Feldman (1988, 2000) offers examples of several obligations which are like this. Some of these (e.g. contractual obligations) are not sufficiently like epistemic ob-

---

⁶ See Feldman (2000, pp. 673-674) for an assessment of this view.
ligations to make the point, while others (e.g., obligations to meet performative norms when engaging in certain roles) are. Thus, Feldman concludes (2000, p. 676) that “It is our plight to be believers. We ought to do it right. It doesn’t matter that in some cases we’re unable to do so.”

Yet another line of response to the voluntarism argument is to retreat from belief to the view that acceptance, which is clearly voluntary, should be the locus of argument evaluation. Lehrer (1981, pp. 79-80; cf. Feldman 1988, p. 240) suggests this route, writing: “Sometimes a person cannot decide what to believe at a moment, but can decide what to accept. … Believing is not an action. Accepting is.” While this escapes the immediate problem raised by the voluntarism argument, it does so at the cost of providing meaningful intentional explanations of actions. That is to say, it gives up Castelfranchi’s (1996, p. 235) Autonomous Cognitive Agent Postulate, and in doing so it fails to adequately explain the intentional behavior of agents.

So, it would seem that there are a series of important conceptual, ontological and empirical questions at the bottom of this debate. The question of whether our beliefs are properly subject to reactive attitudes, seems to depend either on the degree of voluntary control required to properly hold believers responsible for their beliefs, or on whether our having an obligation entails that we can meet it.

9. BELIEF REVISION, ARGUMENTATION AND REASON TRACKING

A second problem arising from the psychology of belief concerns our competence as reason trackers. Belief-centric theories of argumentation hold that argument involves inference, and that rational persuasion occurs when arguers base their cognitive attitudes on reasons. Thus, patterns of inference map out, or represent, episodes in the cognitive lives of reasoners.

As Pinto (2001, p. 10) writes, persuasive argumentation can usefully be understood as “the attempt to modify conscious attitudes through rational means.” Pinto (2001, pp. 37, 32) describes arguments as “invitations to inference” where inference is explained as “the mental act or event in which a person draws a conclusion from premises, or arrives at a conclusion on the basis of the consideration of a body of evidence.” Similarly, Johnson (2000, p. 150) describes persuasive argument as “discourse directed toward rational persuasion,” where rational persuasion is taken to mean “that the arguer wishes to persuade the Other to accept the conclusion on the basis of the reasons and considerations cited [in the argument] and those alone.” Yet, previously Johnson (2000, p. 24) introduced the term “inference” to mean “a movement (of the mind) from one item (usually a thought) to another, where the former serves as the basis for and leads to the later.” These accounts do not differ significantly from Castelfranchi’s (1996, p. 238) conception of inference as a “knowledge acquisition process,” “by which a cognitive system is able to generate, internally, new pieces of knowledge from already existing pieces of knowledge explicitly represented.” (Here, the normative aspect of inference is implicit in its ability to generate knowledge.) Thus, in order for the reasons
cited in the argument to provide the genuine bases for an arguer’s acceptance of the conclusion, the basing relations of an argument’s premises and conclusions must accurately represent the basing relations between reasons and claims in the arguer’s beliefs. That is, arguments must represent and occasion (or bring about) the corresponding and actual inferences which arguers make.

Arguments, on this view, contain or represent inferences – i.e., mental acts on the part of their audiences. On the traditional picture, inference has a dual-nature. On the one hand, inference is a mental process – that is, it is an episode in the psychological history of a cognitive agent. Indeed, as Harman (1986, p. 207) points out, “We normally think of inference as a causal process, in which the premises are all causally operative in producing a new belief (or other doxastic change).” But, not just any kind of mental act, or causal movement of the mind from one idea to another, will qualify as an inference. Purely associationistic causal relations of ideas will not count. Rather, to count as an inference, some thoughts must serve as the basis of others, from which the latter are taken to follow. These basing relations are intrinsic and distinguishing features of inference as a mental process. Thus, Johnson (2000, p. 94; cf. 98) describes inference as “the transition of the mind from one proposition to another in accordance with some principle.” The principle articulates the basing relation that is taken to hold between reasons and conclusions and which justifies or warrants the inferential move. Inference, then, seems to have a dual character on which it is both essentially psychological and causal and on the other hand inherently normative and justificatory. Thus, it would seem that arguers and reasoners must track basing relations among their beliefs in order to infer successfully.

This picture of the nature of inference and its relationship to argument has important consequences for which theories of justification will count as acceptable. In Change in View (1986, pp. 29 ff.) Harman considers two models of ideal belief revision: the foundations theory and the coherence theory. These models are distinguished according to the theory of justification each adopts, and the key feature is whether rationality requires tracking reasons (or the basing-relations among beliefs). Foundationalists say “yes” while coherentists say “no.” As Harman (p. 30) writes, “the theories are most easily distinguished by the conflicting advice they occasionally give concerning whether one should give up a belief P ... when P’s original justification has to be abandoned.” On the foundational view justification is requires reason-tracking and to the extent that reasons are not tracked – let alone not cogent – beliefs become irrationally held. By contrast, the coherence view rejects reason-tracking in favor of a kind of principle of conservatism whereby merely having a belief counts as a reason entitling one to keep it, so long as there is no special reason for abandoning it. It would seem that the inferentialist picture described above is committed to something like the foundationalist picture of rationality.

Paglieri and Castelfranchi (2006b, p. 362) write that one of the problems with all coherence theories of belief revision, such as the AGM model, is that they
do not take into account such ‘reasons to believe’ – and that is precisely the reason why they cannot capture argumentation structures effectively. … Argumentation theories capture the ways in which a desired change in the audience’s beliefs is brought about by the arguer: therefore, without an explicit theory of the reasons to believe something, the main point of argumentation is lost.

Indeed, in previous work Castelfranchi (1996, p. 238) had already postulated the *trace hypothesis* that “we maintain in our memory a trace of the derivation of the cognitive item: its story. … A trace means a link, a relation between the source of the knowledge acquisition or elaboration, and the result. This link maintains also the kind of derivation: perception, communication, reasoning.” The satisfaction of this trace hypothesis, Castelfranchi (*ibid.*.) takes to be an “important cognitive property” of “any process of knowledge acquisition.”

Yet, as crucial as this idea that reasoners and arguers are reason-trackers is to standard (indeed, one might say, informative) accounts of inference and argument, there is reason to think that it is empirically false. As, Harman (1986) observes, when we consider actual cases of belief perseverance in the face of the manifest undermining of their original evidential basis, as experimentally demonstrated in Ross, Lepper and Hubbard’s (1975) debriefing paradigm, it becomes clear that our actual and typical behavior presents significant problems for foundationalist theories.\(^7\) It is found that people regularly retain beliefs even after the positive and manifest refutation of all the evidence upon which their beliefs were originally based. That is, studies such as the debriefing paradigm seem to show that reasoners aren’t reason-trackers. Yet, if this is the case, then it seems that the foundational theory fails by *reductio*. Harman (1986, p. 39) writes, “since people rarely keep track of their reasons, the [foundational] theory implies that people are unjustified in almost all their beliefs. This is an absurd result” (39). Harman’s critique not only challenges the causal elements of the traditional picture whereby beliefs are established on the basis of the reasons supplied in argument, but it challenges the very picture of rationality on which argumentation theory may depend.

10. CONCLUSION

This chapter has sought to recognize contributions made to the theory of argumentation by the work of Castelfranchi, together with Pagliari, by situating their efforts in the context of recent developments in argumentation studies market by the shift in focus from argument to argumentation. Within the last generation it has been realized that in order to properly understand arguments they must be studied not as decontextualized, abstract objects, but rather in the context of their use and application in concrete situations by social and cognitive beings. This process-based approach conceives of argument as an activity, deliberately and often strate-

\(^7\) See Godden (2008) for an extended discussion of this problem.
ically engaged in by intentional agents seeking to affect each other’s behavior by employing reasons to change each other’s minds. It was argued that the process-based view of argument entails the necessity of a psychological dimension in our theories of it. Explanations of intentional, argumentative behavior as well as the inherently cognitive aspects of arguing understood as involving inference and reasoned change in view require a belief-based approach to the theory of argumentation as an activity, as well as the structure of argument as an artifact.

The work of Castelfranchi, together with Paglieri, has contributed substantially to these ends. Pioneering work by Castelfranchi set forth an ontology of beliefs in relation to other mental states such as goals and intentions such that the connection between belief and action can be better explained and understood. With this, the argumentative behavior of rational agents of can be better understood and explained. Further, the Data-oriented model of Belief Revision proposed by Paglieri and Castelfranchi provides theorists with a robust model of rational belief change in relation to the processes of information management which precede and inform belief revision. The DBR model recognizes these processes to be effectively independent, though data management processes can result in inputs, through belief selection, to the processes of belief revision such as inference and argument. In addition to being able to represent such well-established models of argument as Toulmin’s Data-Warrant-Claim model, DBR promises substantial contributions to the theory of argument in ways largely unexplored. For example by formalizing the process of belief selection, it is able to provide an explanation of how and why arguers make determinations of premise adequacy.

There are, though, a variety of problems facing all belief-based approach to argument, many of which come from the psychology of belief. Two of these were raised in the paper.

(i) If beliefs are properly subject to the praise and blame of our reactive attitudes, then it seems as though rational believers must have sufficient control over their beliefs such that they can be rightly held accountable for them. Yet, arguably, believers do not have sufficient voluntary control over their beliefs that they should be held responsible for them. Thus, it seems as though belief is not properly subject to rational appraisal.

(ii) Foundationalist accounts of rationality commonly employed in argumentation theory presuppose that arguers track reasons on the basis of which they draw the conclusions of their inferences and arguments. Yet, empirical studies tend to suggest that reasoners are not reason-trackers in anything like the manner or degree required by such foundationalist accounts of rationality. If so, then it would seem as though either we are much less rational than we ordinarily take ourselves to be, or foundationalism is the wrong picture of rationality to apply to belief revision and argument.
By failing to focus on the inherently doxastic dimensions of argumentation, social, process-based accounts were able to avoid addressing pressing concerns such as these. Yet, when the relationship between belief and argument is articulated, problems such as these come to light – problems which seem to challenge many of the core assumptions on which standard theories of rational belief change, and their attendant accounts of argumentation are based. These problems must be faced squarely, since a retreat to a view of argument which fails to incorporate belief is neither plausible nor viable.

No solution to these problems is here offered. Rather, I raise them merely because they are troubling to me – to my own projects – and it seems to me that they are shared by any theory which seeks to adequately describe, explain and evaluate what is going on when people argue. The turn towards the humanized, process-oriented, situated and concrete view of argument opens not only the doxastic door, but the empirical one as well. Theories hoping to succeed at this level, whether in projects of a descriptive and explanatory nature, or in the normative and evaluative ones with which I am primarily concerned, must aim at a much higher standard. Purely social and behavioral approaches contented themselves to ‘work outside the black box’ of belief and cognition, and were the worse-off for it. Yet, once the ‘black box of belief’ the picture becomes not only more informative but more complicated and challenging.

REFERENCES


Pinto, R.C. (2003). Reasons. In F.H. van Eemeren et al. (Eds.), *Anyone who has a view: Theoretical contributions to the study of argumentation* (pp. 3-16). Dordrecht: Kluwer.


