

## Chapter 5

# Pushing the Bounds of Rationality: Argumentation and Extended Cognition

*David Godden*

Michigan State University, East Lansing, Michigan, USA, dgodden@msu.edu

**Abstract.** One of the central tasks of a theory of argumentation is to supply a theory of appraisal: a set of standards and norms according to which argumentation, and the reasoning involved in it, is properly evaluated. In their most general form, these can be understood as rational norms, where the core idea of rationality is that we rightly respond to reasons by according the credence we attach to our doxastic and conversational commitments with the probative strength of the reasons we have for them. Certain kinds of rational failings are so because they are manifestly illogical—for example, maintaining overtly contradictory commitments, violating deductive closure by refusing to accept the logical consequences of one’s present commitments, or failing to track basing relations by not updating one’s commitments in view of new, defeating information. Yet, according to the internal and empirical critiques, logic and probability theory fail to supply a fit set of norms for human reasoning and argument. Particularly, theories of bounded rationality have put pressure on argumentation theory to lower the normative standards of rationality for reasoners and arguers on the grounds that we are bounded, finite, and fallible agents incapable of meeting idealized standards. This paper explores the idea that argumentation, as a set of practices, together with the procedures and technologies of argumentation theory, is able to extend cognition such that we are better able to meet these idealized logical standards, thereby extending our responsibilities to adhere to idealized rational norms.

## 1. Logic and Reasoning: The Prescriptivity Gap

### *1.1. The Standard Picture and the Path to Prescriptivity*

Logic, Frege tells us in his 1918 essay “The Thought”, describes the laws of truth, from which prescriptions for asserting, thinking, judging, and inferring follow. Until recently, logic’s ‘path to prescriptivity’ for reasoning, and thereby argument, was taken to be relatively straightforward. (Let us call the correct, prescriptive norms for thinking (i.e., reasoning and inference), judgment, assertion, and argument, whatever these turn out to be, ‘rational norms,’ or ‘norms of rationality’). According to the *standard picture* (Stein, 1996, p. 4), logic provides rational norms.<sup>1</sup>

---

<sup>1</sup>The norm of rationality can be applied both to reasoners, when evaluating their rational doings (e.g., changes in view, inferences, and arguings), and to items in their ‘web of belief’ (e.g., propositional attitudes), when evaluating the cogency of the reasons on the basis of which the attitude is held. As such claims can be rationally held when doing so is permitted by the norm of rationality, just as changes in view can be made rationally when

According to this picture, to be rational is to reason in accordance with principles of reasoning that are based on rules of logic, probability theory and so forth. If the standard picture of reasoning is right, principles of reasoning that are based on such rules are normative principles of reasoning, namely they are principles we ought to reason in accordance with.

Today, by contrast, the prevailing view is that the path from logic to rational norms is, at best, indirect.

## ***1.2. The Prescriptivity Gap***

As Harman (2002, p. 171; cf. 1984, 1986, 1995) states, “Inference and implication are very different things and the relation between them is rather obscure.” For instance, adapting Harman’s (1995, p. 184) example, the two sentences:

- (1) Some set of premises, P, implies some conclusion, C, and
- (2) If you believe P you should (or may) infer (or conclude) C,

say quite different things. As Harman notes (1995, p. 184), firstly, they are about entirely different things: (1) is about implication while (2) is about inference. “Inference and reasoning are psychological processes leading to possible changes in belief (theoretical reasoning) or possible changes in plans and intentions (practical reasoning). Implication is most directly a relation among propositions.”

This difference in subject matter has an important consequence when understanding the descriptive relationship between logic and reasoning. According to Harman, logic does not describe reasoning processes—logic is not properly understood as representing some internal, psychological reasoning process. Considering something he calls the “deductive model of inference,” according to which when we reason rightly we do so in accordance with logical rules, Harman objects that the reasoning processes involved in proof construction are neither accurately nor appropriately described in terms of the logical steps in the constructed proof.

Except in the simplest cases, the best strategy is not to start with the premises, figure out the first intermediate step of the proof, then the second, and so on until the conclusion is reached. Often it is useful to start from the proposition to be proved and work backward. It is also useful to consider what intermediate results might be useful. (2002, p. 178)

The reasoning processes involved in proof construction are a kind of problem solving, whereby one attempts to figure out what sequence of logical rules applied to stated premises or derivable theorems will produce the required conclusion. While this reasoning makes use of logical rules by applying them, it is not accurately described by them, nor does it proceed according to them. “The so-called deductive rules of inference are not rules that you follow in constructing the proof. They are rules that the proof must satisfy in order to be a proof” (Harman, 1995, p. 193).

---

similarly permitted. Derivatively, reasoners can be said to be ‘rational’ when they generally behave (e.g., believe and reason) in ways permitted by the norm of rationality.

Not only does logic lack a descriptive relationship to reasoning, it lacks a straightforwardly prescriptive one as well. As Harman (1995, p. 184) observes, logic, in and of itself, is not prescriptive. For example, (1) does not “say anything normative about what anyone ‘should’ or ‘may’ do.” And indeed, (2) is not a consequence of (1). Supposing (1) to be true, even if someone believed P, they might not thereby be justified in either believing or concluding C without explicitly drawing the inference from P to C and thereby recognizing the warranting relation of consequence obtaining between them. Whereas logical rules constitute relations of consequence among propositions, rational norms are epistemic or prudential in nature, and pertain to warranting or justificatory relations among reasons and claims.

This relationship between logic and reasoning creates a *prescriptivity gap* such that, at the very least, rational norms do not follow directly from standard logic.

### 1.3. Principles of Logic and Rational Norms

This prescriptivity gap is widened when it is asked whether logic provides the basis for fit rational norms at all. Consider, for example, the two basic logical relations of consistency and consequence, and some standard principles of deduction that follow from them. First is the principle of *non-contradiction*: that for any well-formed expression, it and its negation cannot both be true, or that standard logical systems are consistent such that logical falsehoods (contradictions) are excluded. Second is the principle of *deductive closure*: that consistent logical systems are closed under deduction, or that all logical consequences of stated premises and axioms are valid theorems of the logical system. Consider now those same principles interpreted or applied as rational norms:

*Consistency*: one’s beliefs should be mutually consistent;

*Closure*: one’s beliefs should be closed under deduction such that one believes all of the logical consequences of one’s present beliefs.

And, in order to satisfy to those norms, one must also adhere to some rational norm like:

*Tracking*: one should keep track of the logical relations between one’s beliefs, such that one can maintain consistency and closure among them.

Harman (1984, 1986, 1995, 2002) argues that none of these are suitable rational principles. Consider consistency: As before, Harman begins by noting (1995, p. 185) that the principles of deduction say very different things than their rational-norm counterparts. “It is one thing to say, ‘Propositions A, B, C are inconsistent with each other.’ It is quite another to say, ‘It is irrational (or unreasonable) to believe A, B, C.’” And, the deductive principle can be true while the rational norm false (1995, p. 185). For example, A, B, and C might be equally well-supported by one’s evidence and yet inconsistent, leaving one in the position of having no reason to abandon any one of them over the others. Further, Harman notes (1984, p. 109), consistency is not a fit rational principle *for us* since “a rational fallible person ought to believe the at least one of his or her beliefs is false,” and believing this will make our entire belief set inconsistent. It would seem that rational yet fallible agents *ought* to have inconsistent beliefs!

Similarly with closure: It is one thing to say that ‘C is a consequence of A and B,’ and quite another to say that ‘One ought to believe C (merely) because it is a consequence of one’s present beliefs A and B.’ Firstly, upon discovering an unacceptable consequence of

one's present beliefs, it might be more rational to revise one's present beliefs, by giving one of them up, rather than to accept the unacceptable consequence. Moreover, there is an infinitude of trivial (e.g., just apply the rule of disjunction introduction repeatedly) or irrelevant (claims of no practical interest or significance) logical consequences of one's present beliefs; surely it is not rational to occupy oneself with these. Indeed, Harman proposes the rational norm of "clutter avoidance" (1986, p. 12; 1995, p. 186) to preempt such patently irrational (in this case, imprudent) behavior.

Thus far, we have seen that "ordinary rationality requires neither deductive closure nor consistency" (Harman 1995, p. 187). Johnson and Blair identify this line of reasoning as the *internal critique*, and describe it as follows: "formal deductive logic is inadequate as a normative theory of argument, supplying neither necessary nor sufficient conditions for logically good argument" (2002, p. 347).

## 2. Bounded Rationality: Widening the Prescriptivity Gap

### 2.1. *The Empirical Critique*

The internal critique is compounded by the *empirical critique* (Johnson and Blair 2002, p. 351), according to which not only does ordinary rationality not require adherence to norms based solely on logical principles, but ordinary rational agents are incapable of adhering to those norms.

The standard picture, on which logic provides rational norms, is committed to the *ambitious claim*, which Perkins identifies and defines as follows:

By and large, people can, should (in the sense of adaptation), and do reason according to standard logic. (2002, p. 189)

The ambitious claim is built upon the principle that *ought implies can*: that the prescriptive force norms have over us presupposes our ability to meet, satisfy, or adhere to those norms. The ambitious claim simply attributes to us the capacity or ability to adhere to rational norms provided by logic.

Problematically, experimental results in the psychology of reasoning indicate that, in predictable and systematic ways, we do not, by and large, adhere to many of the rational norms provided by logic (Perkins 2002). To cite but two well-known examples, Wason's (1968) selection task for deductive reasoning and Tversky and Kahneman's (1982) representative description task for probabilistic reasoning indicate that otherwise rational human reasoners seemingly fail to reason in logically correct ways. More generally, Godden (2012b) notes that we seem not to be reason trackers in the way required by the standard view. And, if we cannot abide by the rational norm of tracking, it is difficult to understand how we should be expected to adhere to the norms of consistency or closure either. Results such as these seem to indicate that the factual, or descriptive, element of the ambitious claim is false (Perkins 2002).

### 2.2. *The Contrapositive of Ought Implies Can*

The normative consequences of these empirical findings are not insignificant. After all, the contrapositive of the *ought implies can* principle is: *cannot implies ought not*. Thus, evidence that ordinary reasoners cannot do something is evidence that they ought not to do it—i.e., that they are under no obligation to do it, and that they cannot be faulted for not

doing it. And, the line of reasoning continues, what better evidence is there that we cannot do something than that, ordinarily, we don't?

Accepting the contrapositive of *ought implies can* involves at least some commitment to the general idea of *bounded rationality* (Simon, 1957), that the prescriptive force of rational norms derives, at least in part, from our ability, whether ordinarily or in principle, to adhere to those norms. Thus, Baron (1985, p. 11) distinguishes between *normative* and *prescriptive* models according to whether they incorporate a boundedly rational element. "A good prescriptive model takes into account the very [cognitive and situational] constraints on time [and resources], etc., that a[n idealized] normative model is free to ignore." This distinction allows us to ask whether a norm should have prescriptive force over an agent (in some situation).

Minimally, the position of bounded rationality commits us to a psychologism whereby the results of psychology are relevant to setting prescriptive rational norms, or at least to explaining the prescriptivity of rational norms. Yet, the extent of any resulting psychologism can still vary greatly. At one extreme, one might insist that the standards of logic remain the proper standards of rationality against which our reasoning ought to be judged, such that our reasonings may rightly be said to be faulty to the extent that we fail to reason in accordance with logical norms. Yet, it might also be said that, to the extent that we are genuinely incapable of reasoning in accordance with logical norms, we should not be held under any obligation to do so, and nor should we be held blameworthy for failing to do so—even though we would be in a better position, rationally speaking, were we better able to do so. Here, the standards of rationality remain unaffected by psychological considerations, though their binding or prescriptive force over us is mitigated by our cognitive abilities. At the other extreme, by contrast, some (e.g., Pelletier & Elio 2005, p. 20) opt for a complete psychologizing of rational norms for non-deductive, default reasoning. They argue: "deductive reasoning has a 'normative standard' that is 'external' to people whereas default reasoning has no such external normative standard ... Here there is no external standard of correctness other than what people actually infer." On this view, the very nature and content of rational standards is determined, at least partly, by psychological considerations. Whatever the extent of the psychologism we adopt in the end, the effect of the *cannot implies ought not* maxim on the prescriptive force of logical principles and the content of rational norms is significant. To appreciate this point, it is useful to consider an account of argumentative rationality and some of the different assumptions embedded within it.

### 3. Argumentative Rationality

#### 3.1. The Basic Idea: Rightly Responding to Reasons

The basic idea of rationality, as it is found in argumentation, is that of rightly responding to reasons (Godden, 2015). Rational belief and action are based on reasons, such that one's beliefs, actions, and changes in view are properly justified and explained in terms of the reasons one has for them (Brown, 1988). Reasoning and deliberation are cognitive processes or activities in which reasons are considered and acted upon. Rightly responding to reasons involves appraising the strength of one's reasons, and adhering to a standard of *evidence proportionality* whereby one accords the degrees of commitment to one's views with the strength of the evidence one has for them (Pinto, 2006, p. 287). As Siegel (1997, p. 2) claims: "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." Argumentation, in turn, can be understood as an interpersonal, communicative activity of "reasoning together" (Campolo, 2005, pp. 38ff.).

### 3.2. Aspects of Argumentative Rationality

Godden (2015) analyses this picture of argumentative rationality, articulating five rough kinds of assumptions contained within it: (i) the normativity assumption, (ii) deontological assumptions, (iii) structuralist assumptions, (iv) internalist assumptions, and lastly (v) the assumption of reflective stability. It is worthwhile to briefly review these assumptions, as it will be shown that several of them are directly challenged by the picture of bounded rationality discussed in the previous section.

First, the normativity assumption highlights the idea that attributions of rationality are honorific value judgements, rather than descriptive statements, and are made by applying behavioral or observational criteria for a prescriptive (e.g., epistemic) standard that we *ought* to meet even in cases where we *do not*.

The deontological assumptions of accountability, obligation, entitlement, and voluntarism identify a cluster of presuppositions latent in our practices of praising and blaming rational agents for their rational behavior, including their cognitive behavior, and in taking ourselves to be right in doing so (Godden, 2010). By holding rational agents accountable for their rational behavior, we not only presuppose that those agents can act rationally—i.e., that they can accord the credence they place in their views to the strength of the reasons they have for them (the assumption of voluntarism), but we also ascribe rational obligations to those agents, such as the obligation to successfully support their views with reasons when called upon to do so and to surrender those views when they cannot meet this obligation, such that they continue to adhere to a standard of evidence proportionalism (Godden, 2014). Rational entitlements are the permissive counterpart to rational obligations: by satisfying our rational obligations we take ourselves to have demonstrated our entitlement to our views, and thereby to be rationally permitted to hold them and to use them in certain ways (e.g., as premises in further inference).

The structuralist assumptions of basing, causal, rule, and tracking unpack some of the ideas implicit in the claim that rationally held views are based on reasons. The basing assumption reiterates the idea that, when a view is based on reasons, there is an explanatory, as well as a justificatory, relationship between the view and the reasons on which it is based, such that changes in the reasons ought to occasion changes in the view. One rationally holds the views one has *because of* the reasons one has for them, and any decoupling here is an indication that one's espoused reasons are not their *actual* reasons. As such, often the force of reasons is understood causally as well as normatively, at least to the extent that the force of reasons can outweigh any non-rational, psychological forces that might affect one's views. The normative force of reasons is typically expressed in terms of warranting or basing rules which explain not only one's *acceptance* of a claim, but the *acceptability* of the claim itself. Understood this way, reasoning is a rule-governed activity, as distinct from mere psychological processes affecting our mental states and attitudes. The tracking assumption articulates the idea that basing one's views on reasons requires monitoring and keeping some account of the acceptability of reasons and the claims one bases on them.

Both the deontological and structural dimensions of argumentative rationality reveal the internalist assumptions of accessibility and articulability. In order that we can rightly be said to have obligations to base our views on reasons which we keep some track of, we must take some cognizance of those reasons (the accessibility assumption), at least to the

extent that we are able to produce or articulate them on demand—e.g., when challenged to demonstrate our rational entitlement to our views.

Finally, in order to rightly respond to reasons, our rationally adopted views must be reflectively stable: once adopted on the basis of reasons, our views should remain settled, and not change due to irrational or non-rational forces, unless and until they are displaced by the force of some stronger reason.

## 4. Bounded Argumentative Rationality

Having set out a picture of argumentative rationality, let us consider the effects of bounded rationality on it. To do this, it is worthwhile to first recognize some of the similarities and connections between rational principles on the standard view and some aspects of the picture of argumentative rationality just presented.

### 4.1. Logic and Rationality Revisited: Paradigms of Irrationality

Despite the prescriptivity gap, logic still seems to make some essential contribution to the bases of rational norms. Consider, for example, the following paradigms of irrationality:

*Manifest inconsistency*: maintaining recognizably inconsistent ‘local’ beliefs (i.e., some limited set of beliefs, among which is not the belief that one is mistaken about at least one of *them*);

*Manifest denial of closure*: refusing to accept the recognizable logical consequences of one’s beliefs;

*Manifest intransigence*: refusing to update one’s beliefs when confronted with a recognizable failure of one’s reasons.

I take each of these cases to be paradigmatically, and hence uncontroversially, manifestly irrational. Yet, these paradigms of irrationality each seem to be explained, at least partly, by their patent illogicality—i.e., by their manifest violation of, or inconsistency with, logical principles.

Each of these paradigms of irrationality has an important connection with some aspect of the picture of argumentative rationality just described. Most generally and importantly, each case seems to violate the principle of evidence proportionality, and thereby the very idea of rationality: that of responding rightly to reasons. More specifically, each involves the failure of tracking some relation, either logical or evidentiary, among one’s reasons and claims, and so contravenes the tracking assumption. The obligation assumption, that we have an obligation to surrender a view whenever we are unable to successfully support it with reasons, is blatantly violated by manifest intransigence. And, insofar as logical inconsistency is the strongest possible reason against some view and logical consequence is the strongest possible support for some view, manifest inconsistency and manifest denial of closure also violate that aspect of the obligation assumption. The manifest-ness of each of these failures implicates that each case also contravenes the other dimension of the obligation assumption, that we have an obligation to successfully support our views with reasons. Unless the failure to meet this obligation is taken as a sign that one’s espoused reasons are not one’s actual reasons, each case also contravenes the basing, rule and causal assumptions. Manifestly and paradigmatically irrational views are not held on the basis of

reasons, which are connected to claims by rules, and so our holding some irrational view is neither caused by, nor explained in virtue of, any reasons we might have or offer for it. Nor, indeed, is it licensed by any rule or reason. In contravening the obligation assumption, each case also contravenes its counterpart, the entitlement assumption, so long as one is never entitled to manifestly, paradigmatically irrational beliefs of the sorts just listed.

It would seem, then, not only that paradigmatic cases of irrationality violate virtually every aspect of argumentative rationality (which, incidentally, is a welcome result for the proposed account of argumentative rationality), but that the irrationality of the belief or behavior is, in each case, significantly explained by its illogicality.

#### ***4.2. Bounded Rationality and Argumentative Rationality***

What is the effect of bounded rationality on this picture? Harman (1986) argues that the logically-based rational norms of consistency, closure and tracking are not suitable as prescriptive rational norms for ordinary human reasoners, on the grounds that we are normally incapable of satisfying any of them. For example, Harman rightly claims that we are incapable of tracking all of the basing relations, and relations of consistency and consequence, that obtain between each of our beliefs. (Indeed, we are incapable of contemplating, in any occurrent sense, an infinite number of beliefs, let alone tracking all of the relations that occur between them.) As such, we are also incapable of adhering to the norms of consistency and closure, since abiding by these requires tracking.

Of paramount importance for our present purposes is the downward pressure that moves of this sort put on rational norms and on argumentative rationality generally. While Harman's position does not directly license any case of paradigmatic irrationality, it comes remarkably close. Consider, for example, the phenomenon of belief perseverance, where beliefs survive "the total destruction of their original evidential basis" (Ross & Anderson, 1982, p. 149), as empirically reported in the debriefing paradigm (Ross, Lepper, & Hubbard, 1975; see Godden, 2012a for a brief survey of this literature). Belief perseverance readily seems to be a case of manifest intransigence, where a subject refuses to update her beliefs despite a recognizable failure of her reasons, and hence seems to be paradigmatically irrational. Yet, Harman (1986, p. 39; 2002; cf. Godden, 2012a, pp. 61ff.) claims that, contrary to any intuition we might have otherwise, since we cannot be expected to track all of the basing relations among our views (being incapable of doing so), belief perseverance is rational.

Now, if Harman is correct that belief perseverance is rational, then the consequences for argumentative rationality are stark indeed. Not only does this view excuse reasoners of one of their basic obligations under argumentative rationality, that of surrendering a view which cannot sufficiently be supported with reasons, but it violates the principle of evidence proportionality, since, by hypothesis, the subject in the debriefing paradigm has no better reason for her persevering belief than for its contradictory. Thus, and most importantly, by permitting cases of manifest intransigence, such a position undermines any promise that reasoning or argumentation might offer for improving the rationality of one's overall view, since reasoners are excused in cases where they refuse to accept the consequences of evidence that they themselves recognize as defeating their own reasons for some occurrent belief. At this point, the very idea of argumentative rationality as rightly responding to reasons seems to have been lost.

Suffice it to say that, insofar as the preceding diagnosis is correct, my prognosis for rational norms, and argumentative rationality generally, is dire. While it might seem as though lowering our rational norms places rationality better within our grasp, succeeding at this diminished standard might not be an achievement deserving of much praise or even

having much epistemic or probative worth in the end. Elsewhere (Godden, 2015), I have sought to address this circumstance by proposing an activity-based account of reasoning that recontextualizes the relationship between reasoning as a justificatory activity and the psychological processes underlying that activity, such that the assumptions of argumentative rationality can better be retained.

## 5. Extending Cognitive Responsibility

In the remainder of this paper, I explore a different treatment. The proposal I offer here suggests that argumentation and critical thinking, understood as a set of practices, together with the procedures and technologies of argumentation theory, can extend our cognitive abilities such that we are better able to meet logically-based rational norms, and thereby extend our responsibilities to adhere to those norms.

### 5.1. *Extended Cognition*

As originally proposed by Clark and Chalmers (1998) and developed by Clark (2008), the extended mind thesis envisions extending the mind artifactually, such that the usual, instrumentalist, account of the role of technologies in the accomplishment of cognitive tasks by human cognitive agents is replaced by an *active externalism* (1998, p. 8) according to which both human agent and technological artifact, when properly connected through a causal coupling that satisfies the ‘glue and trust’ conditions (Clark, 2010), become constituents of larger cognitive system. The justification for the move to count extra-cranial operations as properly cognitive, and thereby the things effecting those operations as components of cognitive systems, derives from the *parity principle*:

If, as we confront some task, a part of the world functions as a process which, *were it done in the head*, we would have no hesitation in recognizing as part of the cognitive process, then that part of the world *is* (so we claim) part of the cognitive process. (1998, p. 8)

Clearly, such an account conceives of cognitive processes functionally, and capitalizes on the multiple realizability hypothesis of functionalism, whereby the same cognitive process can be realized by very different causal processes. The classic example here is memory-impaired Otto who, by making assiduous use of a notebook which he always keeps ready-to-hand, extends his cranial memory to include those things he records in his notebook. According to Clark and Chalmers (1998, pp. 12ff.), Otto’s notebook, and his use of it in the course of his ordinary activities, is on a par with Inga’s neural memory, and her acts of recollecting: “the essential causal dynamics of the two cases mirror each other precisely.” (For completely innocuous examples, consider our ordinary use of prescription eyeglasses or hearing aids.)

#### 5.1.1. *Socially Extended Cognition*

The idea that minds can be artifactually extended so as to include instruments with which cranial cognizers are causally-coupled has been developed to include other ‘technologies,’ broadly understood to include social practices and institutions. (Indeed, Clark and Chalmers anticipate such developments, speculating in their original article as to the possibility of “socially extended cognition” and a “linguistically-enabled extension of cognition” (1998,

pp. 17, 18). Perhaps the most ambitious version of the extended mind is Gallagher's (2013) *socially extended mind*, which "builds on the enactive idea of social affordances."

Just as a notebook or a hand-held piece of technology may be viewed as affording a way to enhance or extend our mental possibilities, so our encounters with others, especially in the context of various institutional procedures and social practices may offer structures that support and extend our cognitive abilities. (Gallagher, 2013, p. 4)

Thus, Gallagher and Crisafi (2009) propose the idea of *mental institutions*, arguing that social institutions such as museums or legal systems meet the parity principle just as well as instruments like calculators and notebooks (see also Gallagher, 2011, 2013).

If we think that cognition supervenes on the vehicle of the notebook, it seems reasonable to say that it supervenes on the vehicle of the museum—an institution designed for just such purposes. (2009, p. 49)

As with *active* externalism, a central aspect of Gallagher's *enactive* externalism is the claim that cognitive processes are distributed across, or realized by, cognitive systems that include both human and technological components. Drawing on De Jaegher and Di Paolo's (2007) idea of *participatory sense making*, Gallagher (2013, p. 8) proposes that social interaction itself has a "certain autonomy ... [that can] transcend the agent's subjective processes" (cf. De Jaegher, Di Paolo & Gallagher, 2010).

Social interaction and participatory sense making specifically involve patterns of engagement that can acquire their own form of self-organization. [...] Participatory sense making is always shaped by super-individual norms and institutional practices. (Gallagher, 2013, p. 8)

Perhaps the most general and universal mental institution, understood as a kind of participatory sense making, is language itself. Recognizing this, Fusaroli, Gangopadhyay and Tylén (2014, p. 37) propose the *dialogically extended mind*, according to which language extends our individual reasoning capacities when we engage in the intersubjective activity of communicatively reasoning together.

Language enables individuals to coordinate their cognitive processes in evolutionarily unprecedented ways, effectively constituting dialogically extended minds. In the skillful intersubjective engagement of symbolic patterns, human beings rely on each other and on established cultural practices to achieve feats that would otherwise be beyond reach.

### 5.1.2. *Extended Cognition: Ontological and Responsibilist Elements*

The ontological dimension of the extended mind thesis is one of its most controversial aspects (Adams & Aizawa, 2001, 2008, 2010). According to Adams and Aizawa, for example, active and enactive externalisms are guilty of the "coupling-constitution fallacy" (2010, pp. 67ff.), which mistakes (i) the essential and fully-integrated use of some technology (be it artifactual or procedural) by some agent within some activity for (ii) some new agent comprised of the closely-coupled things. As an alternative, Hubner (2013), for example, argues for a view of cognition as socially *embedded*, such that technologies are conceived of as "contextual factors" and "enabling conditions" for extending cognitive *capacities*, rather than as "constitutive elements" of newly-conceived cognitive *systems* as proposed by active and socially enactive *extended* approaches. To highlight this kind of distinction, I will distinguish between the *extended mind* thesis (which I will take to have

the ontological commitments just described) and the *extended cognition* thesis (which I will take to be independent of those ontological commitments).

My interests are in extending our cognition. I seek ways that we can extend our cognitive abilities so that our rational responsibilities can likewise be extended, while still adhering to the *ought implies can* maxim. I suggest that this can be done by employing ideas from the extended cognition literature, without taking on any of its controversial metaphysical baggage. One way to do this might be to explain extended cognition, not as the activity of some awkwardly extended mind, but by considering it as the result of some cognitive agent's use of a *regulative technology*—a technology (be it artifactual or procedural) that is designed or used to aid some cognitive agent in the normative regulation of their, or another's, behavior (including their cognitive behavior).

A second controversial issue arising from unqualified or uncritical versions of extended cognition is “cognitive outsourcing” (Menary, 2012). The problem here is one of assigning credit—i.e., praise or blame—for any cognitive, including rational, accomplishment or failure in cases where cognition is extended. Consider, for example, Gallagher's (2013, p. 5) claim that “I cannot remember where the restaurant is, but I, *plus* my technology, can” (Gallagher, 2013, p. 5). It is difficult to accept that the tourist with a GPS-enabled mapping app on his smartphone has the same knowledge and cognitive abilities as the lifetime London cabbie, certified with ‘the knowledge.’ Clearly, the credit for their respective navigational abilities (even supposing them to be functionally equivalent) due to the cabbie vastly exceeds that due to the tourist. (Here, one might even imagine the passenger in the cab, whose only role in the cognitive system, which though trivial is essential nevertheless, is to state her destination to the cabbie.)

One solution to this problem has been to incorporate a virtuous element into the picture of extended cognition such that “cognitive processes that extend outside of the skin of [the] agent can count as part of one's cognitive agency just so long as they are appropriately integrated within one's cognitive character” (Pritchard, 2010, p. 145). Similarly, Roberts (2012, p. 133) argues that “true cognitive extension occurs only when the subject takes responsibility for the contribution made by the non-neural resource.”

One account that strikes a promising balance on these points is Menary's (2007, 2012, 2013) *integrationist account* of extended cognition. On an integrationist account cognition is extended through enculturation: “cognitive capacities are extended through socio-cultural practices” called *cognitive practices* (2013, pp. 26, 29ff).

The practices are patterns of activity spread out across a population. So for example mathematical practices, such as the partial products algorithm, extend the basic biological capacities with which we are endowed. The practice is first learned by manipulating symbols on a page (for example) and becomes a capacity that can be enacted either by bodily manipulation of public symbols, or offline simulations of such manipulations. (2013, p. 26)

A central feature of cognitive practices is that they are essentially normative (2013, p. 29). That is, they can be performed correctly or incorrectly, are situated in other normative activities such as teaching and correction, and are acquired through practice and training. As such, an integrationist account of extended cognition is a good fit with a view of cognition as a rule-governed activity rather than as a psychological process (cf. Godden, 2015). Further, “[m]any of these [cognitive] practices involve artifacts such as tools, writing systems, number systems, and other kinds of representational systems” (2013, p. 29). Yet, the role of artifacts in cognitive practices is not that of a constituent in some larger cognitive system, but that of a regulative technology. In this way, an integrationist account avoids the cognitive outsourcing of strictly causal, artifactual accounts, and instead

incorporates the normative and deontological elements demanded not only by credit-based, virtue accounts of knowledge and cognitive achievement, but by the picture of argumentative rationality presented in section 3.

## ***5.2. From Extended Cognition to Extended Responsibility***

Importantly, by providing an inherently normative extension of our cognitive abilities, cognitive practices thereby extend our cognitive responsibilities. To see this, consider an example adapted from Menary (2013, pp. 29ff.). Suppose I have an arithmetic test where I am expected to correctly multiply large numbers together. Should I be excused from having to complete the test, or for failing it, on the grounds that the numbers are so impossibly large that I could not conceivably determine the product in my head using only my untutored arithmetical intuitions? Well, no. Instead, I can be expected to learn and apply the partial products algorithm in order to calculate the final product, even if I cannot compute it in my head all at once.

Similarly, suppose it is my job to track a baseball game. That is, I am to make note of, and subsequently report on, every pitch and every play of the game. Such a feat is typically well beyond the normal cognitive abilities of the ordinary baseball spectator. How, then, can some people be obliged to do this as part of their job? Well, quite easily actually. One makes use of a “scorecard”: a table on which one records in writing how each pitch and play is called, as a means of fulfilling one’s obligation to note and report on the details of the game. Eventually, one might even internalize this practice such that one can accomplish it without having to write it down in pen-on-paper, but instead develop, through practice, the habit of making a mental note of each play such that one can recall, at least for a time, an entire game in one’s head.

The initial point here is that there are ordinary cases where one’s constitutive yet unaided inability to perform some action does not relieve one from the responsibility of doing so. Oftentimes, there are readily available cognitive practices or regulative technologies (whether instruments or procedures) whose very purpose is to aid one in meeting their obligations in such circumstances. Particularly when significant social value is placed upon either the obligations themselves or the ends to which the obligations contribute, those practices and technologies are frequently instituted as facets of the very activities or social practices in which one is engaged. Thus, rather than be excused from meeting one’s obligations in such circumstances, one is instead obliged to avail oneself of the available or instituted cognitive practices or regulative technologies in order that one *can* meet those obligations.

## **6. Extending Cognition Through Argumentation**

Having proposed the idea that cognitive practices and regulative technologies can extend our cognitive responsibilities by extending our cognitive abilities, let me conclude by suggesting some of the ways that the resources of argumentation theory serve to extend our cognitive abilities and thereby extend our rational responsibilities.

### ***6.1. Argumentative Resources for Extending Rational Responsibilities***

As noted earlier, the internal and empirical critiques present a significant and direct challenge not only to logically-based rational norms such as consistency, closure and tracking, but also to many of the core elements of argumentative rationality. At least one

element of the empirical critique is the claim that we are not sufficiently able reason-trackers that we should be expected to avoid manifest intransigence, as evidenced by the diagnosis of belief perseverance as rational. Happily, argumentation theory provides several resources that can significantly extend our reason-tracking abilities.

### *6.1.1 Argument Diagramming or Mapping*

First among these are the techniques and technologies of argument mapping or diagramming. These come in all varieties, from simple pen-on-paper practices of argument analysis, whereby the premises, conclusions and basic patterns of inference are identified and itemized, to fully automated and scalable mapping software, allowing several agents to collectively construct vast networks of nested claims and reasons where inferential patterns, schemes, required-but-unstated premises and critical questions or potential defeaters are automatically supplied. Basically, argument diagrams work as ‘argumentative scorecards,’ allowing reasoners to better track the different commitments and claim-reason complexes that are ‘on the table’ in some given argumentative exchange. Even the technique of Venn diagramming allows some reasoners to recognize cases of valid consequence, invalidity, consistency, and inconsistency in ways that are not wholly apparent to them when a putative syllogism or immediate inference is presented textually or verbally. More robust and versatile mapping technologies facilitate the efficient and effective tracking of a multitude of rationally-significant relations between claims (e.g., consequence, consistency, closure, exclusion, likelihood, evidence, support, relevance, dependency, defeat, coherence, explanation, etc.).

### *6.1.2 Procedural Norms for Argumentation*

A second cluster of resources developed by argumentation theorists stems from procedural approaches to reasoning and argumentation. Again, these come in a wide variety ranging from rough-hewn Waltonian dialogue types, to the Pragma-Dialectical model of a critical discussion, to fully operationalized, program-like rule systems. Procedural approaches seek to provide the ‘partial-product algorithms of reasoning.’ In their most rigorously articulated versions, they furnish rule-governed, step-by-step procedures that reasoners can follow with relative ease in order to reach a rational resolution to a difference of opinion or to settle the rational acceptability of a standpoint. Here again, the tasks of tracking, and maintaining consistency and closure are internalized into the procedural rules themselves such that merely following the rules at any given point in the process ensures that one meets the rational norms embodied within the model.

### *6.1.3 Schemes and Fallacies: Commonplace Recipes for Improved Rationality*

A less comprehensive but more perhaps more easily acquired and employed set of argumentative devices are the argumentation schemes and their accompanying critical questions (Walton, Reed, & Macagno, 2008), and fallacies. These function more as rational topoi, or commonplaces, and can be seen as the ‘on-the-box recipes of reasoning.’ By using these devices, one needn’t know how to construct or critique an argument from scratch, so long as they can follow directions by assembling and arranging the listed ingredients. Argumentation schemes provide common patterns of cogent yet defeasible presumptive reasoning, each of which can then be evaluated for stereotypical points of defeat by

applying the attendant critical questions (each of which is designed to evaluate some aspect of cogency (Godden & Walton, 2007)). The counterparts to schemes are fallacies, which catalogue typical patterns of incogent reasoning that are frequently but mistakenly appraised as cogent. Because these devices are not comprehensive, even their judicious employment cannot guarantee adherence to every rational norm. That said, they are clearly capable of improving one's rational abilities and frequently produce effective, satisfactory results. And, to the extent that they succeed in these respects, they also extend our rational accountability and accomplishments.

#### *6.1.4 Guidance Norms: From Lists of Do's and Don'ts to Proof Systems*

A final cluster of resources provided by argumentation theory—drawing upon the results of adjacent fields like critical thinking, epistemology, and logic—is designed for the task of determining, or estimating, and critically evaluating, the probative force of reasons. As with the tools and techniques already discussed, these evaluative technologies exhibit varying degrees of precision, scope, and rigor (or formalization). At one extreme, there are collections of commonplace rules that provide good starting places in our projects of rational evaluation. For example, Feldman's (1999) "Basic Rules of Argument Evaluation" offers a list of Do's and Don'ts for critically appraising premise acceptability and inferential strength, such as: "direct criticisms at individual premises," "don't criticize an argument by denying its conclusion," and "make your criticisms substantial," "don't accept competing arguments." An equally informal, but completely generic technique for appraising inferential strength is provided by the method for constructing counterexamples: describe a possible situation (or, the most plausible situation) in which an argument's premises can all be true while its conclusion false; evaluate the strength of the inference by comparing the relative likelihoods of the counterexample with the truth of the conclusion given the premises. More thorough, though fit for a narrower range of (specifically causal) inferences, are Mill's (1973, pp. 388ff) methods of experimental inquiry (or canons of induction): the methods of agreement, difference, joint method of agreement and difference, residues, and concomitant variation. These methods aid not only hypothesis formation, but also the testing of formulated hypotheses. Methods like these can be operationalized with varying degrees of rigor, completeness, and comprehensiveness that extend well beyond ad-hoc lists of instructions. For example, Flage (2000) provides a set of flowcharts for critical thinking that map out, in step-by-step fashion, a "highly structured decision-procedure" for the evaluation of argument, and include subsidiary decision trees for dealing with things like ambiguity, relevance, presumption, and observation, testimony and surveys. A final technology, of course, is found in the formal proof systems of formal logic and the probability calculus. As with the procedural rules of some dialogue systems discussed above, these logical systems embed the tasks of tracking, and maintaining consistency and closure, within the very structure of the system itself, by fully operationalizing basic norms such as consistency and consequence.

#### *6.2. Closing Remarks: The Virtues of a 'Can Do' Attitude*

While these are but four samples of the kinds of wares that argumentation theory has to offer, the more general point should be readily apparent. Argumentation theory has designed, constructed, and sometimes imported a diverse and versatile product line of turnkey cognitive practices and off-the-shelf regulative technologies that are remarkably effective at extending our untutored, unaided, intuitive rational abilities. In view of their ready availability, the "cannot, because does not, therefore ought not" line of reasoning

frequently advanced by theorists of bounded rationality, together with its overtly psychologistic counterpart “ought because does,” should be received with considerably more skepticism and reservation than has been fashionable of late.

Overall, argumentation theory informs a set of mental institutions into which human reasoners can readily be enculturated. And, this enculturation into the social practices of critical reasoning and argumentation is perhaps the most basic, yet most important, result of the cognitive practices and regulative technologies offered by argumentation theory. That people are willing to hold themselves and those around them rationally accountable, such that they see the value in, and are willing to take on for themselves, the basic rational obligations of giving reasons for their views and changing those views when their reasons don't pan out, is the kernel of rationality understood as rightly responding to reasons. Having this sense of rational accountability makes one a scorekeeper in the game of giving and asking for reasons (to borrow Brandom's (1994) phrase), and provides the impetus to seek out and develop the skills required to rightly respond to reasons in increasingly refined and effective ways.

Those of us who feel committed to, or even find something right about, the picture of argumentative rationality sketched above, or to the notion that at least some rational norms derive at least partly from logical ideals, should take heart at the ‘can do’ attitude embodied by critical thinkers, of every ability, everywhere.

## References

- Adams, F. & Aizawa, K. (2001). The bounds of cognition. *Philosophical Psychology*, *14*, 43-64.
- Adams, F. & Aizawa, K. (2008). *The bounds of cognition*. Oxford: Blackwell.
- Adams, F. & Aizawa, K. (2010). Defending the bounds of cognition. In R. Menary (Ed.), *The extended mind* (pp. 67-80). Cambridge, MA: MIT Press.
- Baron, J. (1985). *Rationality and intelligence*. Cambridge: Cambridge University Press.
- Brandom, R. (1994). *Making it explicit: Reasoning, representing, and discursive commitment*. Cambridge, MA: Harvard University Press.
- Brown, H. (1988). *Rationality*. London: Routledge.
- Campolo, C. (2005). Treacherous ascents: On seeking common ground for conflict resolution. *Informal Logic*, *25*, 37-50.
- Clark, A. (2008). *Supersizing the mind: Reflections on embodiment, action, and cognitive extension*. Oxford: Oxford University Press.
- Clark, A. (2010). Memento's revenge: The extended mind, extended. In R. Menary (Ed.), *The extended mind* (pp. 43-66). Cambridge, MA: MIT Press.
- Clark, A. & Chalmers, D. (1998). The extended mind. *Analysis*, *58*, 7-19. [Reprinted in Menary, R. (Ed.). (2010). *The extended mind* (pp. 27-42). Cambridge, MA: MIT Press.]
- De Jaegher, H., & Di Paolo, E. (2007). Participatory sense making: An enactive approach to social cognition. *Phenomenology and the Cognitive Sciences*, *6*, 485-507.
- De Jaegher, H., Di Paolo, E., & Gallagher, S. (2010). Does social interaction constitute social cognition? *Trends in the Cognitive Sciences*, *14*, 441-447.
- Feldman, R. (1999). *Reason and argument* (2nd ed). Upper Saddle River, NJ: Prentice Hall.
- Flage, D. (2000). Flowcharts for critical thinking. *Informal Logic: Teaching Supplement #3*, *20*, TS57-TS69.
- Frege, G. ([1918] 1956). The thought: A logical inquiry. *Mind*, *65*, 289-311.

- Fusaroli, R., Gangopadhyay, N., & Tylén, K. (2014). The dialogically extended mind: Language as skillful intersubjective engagement. *Cognitive Systems Research*, 29-30, 31-39.
- Gallagher, S. (2011). The overextended mind. *Versus: Quaderni di Studi Semiotici*, 55-66.
- Gallagher, S. (2013). The socially extended mind. *Cognitive Systems Research*, 25-6, 4-12.
- Gallagher, S. & Crisafi, A. (2009). Mental institutions. *Topoi*, 28, 45-51.
- Godden, D. (2010). The importance of belief in argumentation: Belief, commitment and the effective resolution of a difference of opinion. *Synthese*, 172, 397-414.
- Godden, D. (2012a). Rethinking the debriefing paradigm: The rationality of belief perseverance. *Logos & Episteme*, 3, 51-74.
- Godden, D. (2012b). The role of mental states in argumentation: Two problems for rationality from the psychology of belief. In F. Paglieri, L. Tummolini, R. Falcone & M. Miceli (Eds.), *The goals of cognition: Essays in honor of Cristiano Castelfranchi* (pp. 123-143). London: College Publications.
- Godden, D. (2014). Teaching rational entitlement and responsibility: A Socratic exercise. *Informal Logic, Teaching Supplement*, 34, 124-151.
- Godden, D. (2015). Argumentation, rationality, and psychology of reasoning. *Informal Logic*, 35, 135-166.
- Godden, D. & Walton, D. (2007). Advances in the theory of argumentation schemes and critical questions. *Informal Logic*, 27, 267-292
- Harman, G. (1984). Logic and reasoning. *Synthese*, 60, 107-127.
- Harman, G. (1986). *Change in view: Principles of reasoning*. Cambridge, MA: MIT Press.
- Harman, G. (1995). Rationality. In E. Smith, & D. Osherson (Eds.), *Thinking: An invitation to cognitive science, Vol. 3, 2<sup>nd</sup> ed*, (pp. 175-211). Cambridge, MA: MIT Press.
- Harman, G. (2002). Internal critique: Logic is not a theory of reasoning and a theory of reasoning is not logic. In D. Gabbay, R. Johnson, H. Ohlbach, & J. Woods (Eds.), *Handbook of the logic of argument and inference: Turn towards the practical* (pp. 171-186). Amsterdam: Elsevier.
- Huebner, B. (2013). Socially embedded cognition. *Cognitive Systems Research*, 25-26, 13-18.
- Johnson, R. & Blair, J.A. (2002). Informal logic and the reconfiguration of logic. In D. Gabbay, R. Johnson, H. Ohlbach & J. Woods (Eds.), *Handbook of the logic of argument and inference: Turn towards the practical* (pp. 339-396). Amsterdam: Elsevier.
- Menary, R. (2007). *Cognitive integration: Mind and cognition unbounded*. Basingstoke: Palgrave Macmillan.
- Menary, R. (2012). Cognitive practices and cognitive character. *Philosophical Explorations*, 15, 147-164.
- Menary, R. (2013). Cognitive integration, enculturated cognition and the socially extended mind. *Cognitive Systems Research*, 25-26, 26-34.
- Mill, J.S. [1843/1872.] 1973. *A system of logic*, Books I-III: *The collected works of John Stuart Mill*, Vol. 7. J.M. Robinson (ed.). London: Routledge and Kegan Paul.
- Pelletier, F., & Elio, R. (2005). The case for psychologism in default and inheritance reasoning. *Synthese*, 146, 7-35.
- Perkins, D. (2002). Standard logic as a model of reasoning: The empirical critique. In D. Gabbay, R. Johnson, H. Ohlbach & J. Woods (Eds.), *Handbook of the logic of argument and inference: Turn towards the practical* (pp. 187-223). Amsterdam: Elsevier.
- Pinto, R.C. (2006). Evaluating inferences: The nature and role of warrants. *Informal Logic*, 26, 287-317.
- Pritchard, D. (2010). Cognitive ability and the extended cognition thesis. *Synthese*, 175, 133-151.

- Roberts, T. (2012). You do the maths: Rules, extension, and cognitive responsibility. *Philosophical Explorations*, 15, 133-145.
- Ross, L., & Anderson, C. (1982). Shortcomings in the attribution process: On the origins and maintenance of erroneous social assessments. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 129-152). Cambridge: Cambridge UP.
- Ross, L., Lepper, M., & Hubbard, M. (1975). Perseverance in self-perception and social perception: Biased attributional processes in the debriefing paradigm. *Journal of Personality and Social Psychology*, 32, 880-892.
- Siegel, H. (1997). *Rationality redeemed: Further dialogues on an educational ideal*. New York: Routledge.
- Simon, H. (1957). A behavioral model of rational choice. In *Models of man, social and rational: Mathematical essays on rational human behavior in a social setting*. New York: Wiley.
- Stein, E. (1996). *Without good reason: The rationality debate in philosophy and cognitive science*. Oxford: Oxford UP.
- Tversky, A. & Kahneman, D. (1982). Judgements of and by representativeness. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgement under uncertainty: Heuristics and biases* (pp. 84-98). New York: Cambridge UP.
- Walton, D., Reed, C., & Macagno, F. (2008). *Argumentation schemes*. Cambridge: Cambridge UP.
- Wason, P.C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology*, 20, 273-281.