The CDF Collaboration and Argumentation Theory:

The Role of Process in Objective Knowledge

William Rehg and Kent Staley

Department of Philosophy

Saint Louis University

Abstract:

For philosophers of science interested in elucidating the social character of science, an important question concerns the manner and degree to which the objectivity of scientific knowledge is socially constituted. We address this broad question by focusing specifically on philosophical theories of evidence. To get at the social character of evidence, we take an interdisciplinary approach informed by categories from argumentation studies. We then test these categories by exploring their applicability to a case study from high-energy physics. Our central claim is that normative philosophy of science must move beyond abstract theories of justification, confirmation, or evidence conceived *impersonally* and incorporate a theoretical perspective that includes *dialogical* elements, either as adjuncts to impersonal theories of evidence or as intrinsic to the *cogency* of scientific argumentation.[*]

Philosophers and scholars of science from a range of theoretical perspectives have insisted that good science crucially depends on the broader social context in which scientific research is conducted. From realists like Karl Popper to feminist standpoint epistemologists like Sandra Harding, theorists across the spectrum accept the idea that scientific activity constitutes not

---

[*] We are grateful for helpful comments on this paper from Peter Achinstein and from an anonymous referee for *Perspectives on Science*. We presented an earlier version of this paper in a workshop at the Philosophy of Science Association meeting in Austin, Texas in 2004, and we are thankful for helpful comments from our fellow workshop participants, Deborah Tollefsen and Henry Frisch, our session chair, John Forge, and our audience, Allan Franklin. We presented an even earlier version at Saint Louis University and we thank our colleagues there for their insights.

simply an intellectual but also a social process.[1] Still at issue, however, is how and to what degree the social character of science not only provides more or less favorable circumstances for good science, but also *constitutes* the scientific claim to rationally justified, objective knowledge.[2]

In this essay we address this issue by focusing specifically on normative theories of evidence. According to Peter Achinstein (2000, 2001, 3), the standard philosophical approaches to evidence are largely irrelevant to actual scientific practice. One reason for such irrelevance, we surmise, lies in the abstract character of standard philosophy of science, which tends to analyze the concept of evidence apart from the social aspects of inquiry. The question of social context thus poses a challenge for such theories: how, precisely, is evidence related to science as a social process? In response, we claim that philosophical analyses can achieve practical relevance through the interdisciplinary study of scientific activity as a set of *argumentative practices*. Normative philosophy of science must therefore move beyond abstract theories of justification, confirmation, or evidence conceived *impersonally* and incorporate a theoretical perspective that includes *dialogical* elements, either as adjuncts to impersonal theories of evidence or as intrinsic to the *cogency* of scientific argumentation.

For this richer dialogical perspective we look to argumentation studies (or "argumentation theory"). As an interdisciplinary field of research, argumentation studies encompasses a broad

---

[1] Popper's emphasis on critical discussion in science implies that scientific institutions should be open, but he also held that science functions best under conditions of democracy (Popper 1957, 154-55; 1989, chaps. 4, 10), a view echoed by Polanyi (1951) and Merton (1973), among others. For Harding's view, see Harding (1998).

[2] For a range of recent views, see Bloor (1991); Kitcher (1993; 2001); Longino (1990; 2002); Solomon (2001); Kusch (2002).

range of perspectives on argument assessment, which originally stem from classical thought but have lately expanded under the pressure of theoretical innovation in areas of rhetoric, dialectic, informal logic, discourse analysis, literary theory, linguistics, and so on (see Eemeren et al. 1996). Other than studies of inductive logic, much of this work has received only peripheral notice by mainstream philosophers of science.[3]

To link up philosophical theories of evidence with argumentation theory, we make two initial moves. First, we treat scientific inquiry as a process of argumentation that issues in arguments, i.e., textual products of scientific research: papers, articles, conference presentations in which scientists attempt to show that certain methods and observations justify some conclusion. The conclusion can be offered with varying degrees of confidence: for example, as a hypothesis that merits further research, as plausible but not established, as conclusively established, and so on. Second, we use the idea of the *cogency* of an argument as a "boundary concept" that can mediate between different disciplines (cf. Klein 1996). In argumentation studies, "cogency" refers to the goodness or strength of arguments, or to their force or acceptability for an audience (Govier 2005; Eemeren et al. 1996, 32, 155). The idea also calls to mind "sufficiency of support" as a criterion of good arguments (Johnson 2000, 204-06), plausibility logics (Rescher 1976), and differentiated burden of proof standards. In philosophy of science, cogency links up with the history of attempts to define degree of inductive support.

[3] Although much post-Kuhnian philosophy of science might be construed as dialectical in spirit, only a few philosophers have made explicit use of the dialectical tradition to analyze scientific argument (e.g., Rescher, 1977; Pera, 1994). Philosophers have also contributed to the "rhetorical turn" in science studies, but the bulk of this extensive literature, as well as the most influential studies, stem from English, composition, speech communication, literary criticism, and linguistics (see Klein 1996, 66-70; Gross 2006; Harris 1997).

With these initial moves in place, our argument then proceeds by developing their implications for both philosophers and scientists. For philosophers, the argumentative turn leads to a specific way of formulating our leading question concerning the social character of scientific rationality: What role, if any, does the social process of scientific inquiry and argumentation play, not only in generating but also in *constituting* the cogency of good scientific arguments? On the received view of good science, socio-institutional context only affected the pace and direction of inquiry—the generation of cogent arguments but not their constitution. Understood as a theory of argument, objective theories of evidence suggest that cogency rests on something like the "intrinsic" or impersonal merits of evidential arguments, even if actual persuasiveness for an audience involved social and psychological factors of various sorts (sec. 1). In large scientific collaborations, however, the social circumstances of argumentation play a significant role in the construction and assessment of cogent arguments, as we illustrate with an actual case of high-energy particle physics (HEP), in which scientists themselves confronted the question of evidence: the writing of the paper announcing the existence of evidence for the top quark. The history of that argument displays two features that are of special interest, which we call *probation by process* and *heterogeneous consensus* (sec. 2). The complications raised by such cases pose a challenge both for scientists—how they understand their collaborative practice of paper writing—and for philosophers—how they conceive of argumentative cogency (sec. 3).

1 Objectivist Theories of Evidence: From Intrinsic to Impersonal Merits

We begin with objective theories of evidence as they might appear when read from the standpoint of argumentation theory. After describing the models of cogency one might ascribe to logical empiricist definitions of confirmation (1.1.) and Achinstein's realist theory of evidence (1.2), we develop the argumentation-theoretic implications for the leading question (1.3).

(1.1) The heavy focus of pre-Kuhnian philosophy of science on formal logical analysis might foster the idea that the cogency of an argument depends solely on its intrinsic merits. Here the paradigm is the sound deductive argument, which seems to enjoy a kind of intrinsic strength: if the premises are in fact true, then the valid structure of the argument guarantees a true conclusion, invulnerable to any additional information. The cogency of sound deductions is intrinsic in the sense of residing solely in properties of the supporting reasons and their relationship to the conclusion.

Not surprisingly, this analogy to deductive arguments is emphasized in those works on confirmation theory that seem to come nearest to the idea of intrinsic merits. We are thinking in particular here of work in the logical empiricist vein such as that of Carnap (1962) and Hempel and Oppenheim (1945). Carnap insists that statements of degree of confirmation, like statements of logical implication, "can be established by a logical analysis" of the sentences in question (Carnap 1962, 201). Likewise, in Hempel and Oppenheim's account, once two sentences H and E in a given language have been specified, the degree of confirmation of H by E "is completely determined by the formal, or syntactical, structures of the two sentences [H and E] alone" (Hempel and Oppenheim 1945, 113).

Such analytic approaches to evidential support must contend against a wide variety of competing views, one of which we consider below. Even supposing that such an analogy to deduction holds, however, ambiguities immediately arise. Both validity and truth depend on things beyond the argument itself, namely the choice of logical system and the phenomena or reality that (empirical) premises describe. One might defend validity as an intrinsic merit by arguing that the logical system simply defines what count as intrinsic connections between sentences. But truth presents a more difficult problem. On one common understanding, truth is indeed a property of statements (and thus is intrinsic), but it depends on the world as the truth-maker (and thus is extrinsic). By defining the degree of confirmation relative to sentences whose acceptance is a matter of pragmatic convention, logical empiricists in effect dodged this question by relegating the intrinsic merits of arguments to logical validity (e.g., Hempel 2000).

The limitations of an intrinsic-merits model of cogency also become evident when one wants to *apply* a formal inductive logic to a given "knowledge situation." At that point, Carnap maintains, one must satisfy a methodological (rather than logical) rule, namely the "requirement of total evidence," which stipulates that the determination of the degree of confirmation for a given hypothesis be based upon "the total evidence available." This rule becomes relevant precisely when one turns to the practical task of assessing the evidence for a particular hypothesis, and deciding whether to believe it (Carnap 1962, 211). Whether that condition has been met is not *obviously* a strictly intrinsic feature of the argument, although one could perhaps salvage its intrinsic character by an *ad hoc* stipulation that each argument be understood to include an implicit statement that the total evidence requirement has been met. Certainly its assessment cannot, in general, be made by direct inspection of the argument itself.

(1.2) Against this background, Peter Achinstein's attempt to go beyond logical empiricist models of evidence is instructive. Achinstein wants to overcome a number of difficulties with those models, including their failure to explain ongoing research when theories are well confirmed (1991, 323-26; 2001, chap. 3). To account for this feature of scientific practice, Achinstein insists, one must bring in the scientist's concern for the truth of evidence. Consequently, his objective theory of evidence turns on a realist conception of *potential evidence* that is missing in Carnap's model. Whereas Carnap's confirmation relations are formal and syntactically defined, and thus determinable a priori, Achinstein holds that evidence e, if true, is potential evidence for h only if certain empirical conditions are met. As we shall see, this makes his theory of evidence more amenable to a conception of cogency based on the broader idea of impersonal merits, rather than merely intrinsic (logical) merits.[4]

Achinstein (2001, 170) defines potential evidence in terms of three conditions:

(PE) *e* is potential evidence that *h*, given background assumptions *b*, iff

1. *p*(there is an explanatory connection between *h* and *e*/*e*&*b*) > 1/2
2. *e* and *b* are true
3. *e* does not entail *h*.

First, some brief clarifications and comments on this account: The requirement of an explanatory connection is met if (1) *h* explains *e*, (2) *e* explains *h*, or (3) some other statement explains both *e* and *h*. The concept of probability in the first condition in (PE) is an objective concept of epistemic probability, such that $p(h/e) = r$ is to be interpreted as asserting that, given *e*, it is reasonable to degree *r* to believe *h*.[5]

Potential evidence is the central concept in terms of which Achinstein defines three others: subjective evidence, ES (epistemic situation) evidence, and veridical evidence. Achinstein claims that it is *veridical* evidence that scientists seek in performing experiments (Achinstein 2001, 34-

[4] We do not claim that Achinstein intends for his account to be used in this way, but instead wish to pursue the question of whether his theory of evidence could be related to cogency along the lines developed here. One important potential difference between the concept of *evidence* as treated by Achinstein and the concept of argumentative *cogency* is that whether *e* is potential evidence for *h* on Achinstein's view is a matter entirely distinct from whether *e* has been shown to be evidence for *h*. Whether a similar distinction can be drawn for cogency we discuss in section 3.

[5] This interpretation distinguishes Achinstein's account from most views of epistemic probabilities, which interpret them as of degrees of belief.

35), where *e* is veridical evidence that *h* if and only if *e* is potential evidence that *h* and *h* is true.[6] For our purposes, however, what is most important about Achinstein's account is the distinction that he draws between potential or veridical evidence, which are not relativized, and subjective or ES evidence, which are relativized to persons and epistemic situations, respectively. It is potential evidence that forms the core of Achinstein's account, and its unrelativized status suggests that it might form the basis for a concept of cogency that can be described as both objective and impersonal. It will therefore be the focus of what follows.

Crucial to Achinstein's position is a further claim about evidence that seems to point toward a connection with cogency:

(GR) If *e* is potential evidence that *h*, then *e* is a good reason to believe *h*.

Presumably, from the fact that *e* is a good reason to believe *h*, it would follow that an argument citing *e* as a premise and *h* as conclusion is cogent. What more could we ask of an argument than that it provide its audience with good reasons for accepting its conclusion?[7]

Significantly, the concepts on both sides of the conditional (GR) are unrelativized. Achinstein distinguishes between two different kinds of justification for epistemic agents. In one sense, what one is justified in believing, or has good reason to believe, depends on one's epistemic situation. But the concept of *good reason* that Achinstein advocates in his analysis of

---

[6] Achinstein also suggests the possibility of an even stronger concept of veridical evidence that also requires that the explanatory connection between *e*'s being true and *h*'s being true must be actual and not merely highly probable (Achinstein 2001, 174).

[7] In light of the distinction discussed below, Achinstein himself need not accept this apparent connection between potential evidence and argumentative cogency. He could take the position that the notion of good reason that is needed for argumentative cogency is one that is relativized to epistemic situation.

potential (and veridical) evidence is an abstract concept according to which if *e* is a good reason to believe *h*, then it is so for anyone, regardless of her epistemic situation, and in spite of the fact that, given her epistemic situation, she may be (in the first sense of justification) completely unjustified in believing *h*. (See Staley 2005 for a critical discussion.)

When *e* is potential evidence that *h*, then *e*'s being a good reason to believe *h* is determined by "certain physical and/or mathematical facts or states of affairs [that] make it reasonable . . . to believe" *h* (ibid., 97). The "normative fact" of good reasons, or cogency, in effect supervenes on such physical and mathematical facts, including the facts expressed by *e* and *b*, as well as those that ground the satisfaction of the "explanatory connection" requirement (expressed in clause 1 of PE above). This supervenience does not depend on the beliefs of the audience or anyone else, or on any epistemic situation, actual or possible.

This metaphysical, knower-independent relation between evidence and facts means that evidence claims are generally potentially "empirically incomplete." That is, presented with a description of an experimental outcome offered in support of some hypothesis, one might need additional information in order to determine whether that outcome is genuinely evidence in support of that hypothesis. This gives evidence claims an empirical, rather than a priori, status. As a result, we cannot plausibly understand Achinstein's theory simply in terms of intrinsic merits. Whereas on Carnap's view confirmation relations can always in principle be "read off" of a complete statement of the reasons offered in confirmation of a hypothesis, Achinstein's theory explicitly denies that this is the case. There is the potential that one may have to look beyond the text presenting the evidential argument, to empirical facts "in the world" (and not considered in the argument) in order to assess the cogency of that argument.

Nor can a notion of intrinsic cogency be reconciled with Achinstein's account by the *ad hoc* move of including an implicit claim to the effect that "there are no interferences at work here beyond those we have taken into account." The kind of additional empirical information that might be needed to assess an evidence claim is not merely the determination of the truth of such a claim, but could include information having to do with data collection procedures, sample size, stopping rules, the adequacy of an experimental set-up, etc. Absent the specification of such information, we may be unable to determine the evidential status of an experimental outcome. Neither can one establish prior to testing what kinds of information must be included in a description of an experimental result in order for the result to have a determinate evidential status as described, though careful pre-trial planning can help to mitigate the difficulties (Staley 2004a, ch. 6).

These differences notwithstanding, both the logical empiricists and Achinstein define cogency in relation to objective properties of the supporting evidence and its relationship to the hypothesis-conclusion, considered apart from the social process of arguing and its pragmatic contingencies. Both intrinsic logical merits and the truth of premises have an impersonal character: they are not relativized to particular persons or audiences (Achinstein's ES-evidence is also impersonal in this sense). We might thus regard these theories of evidence as defining cogency in terms of *impersonal merits*.

(1.3) To develop this point further in terms of argumentation theory, we can say that the above-described approaches to evidence provide two different ways of understanding cogency as based entirely on the impersonal merits of the argument as a "product." Argumentation theorists typically approach arguments as the *products* of argumentative practices, understood as a type of social practice (e.g., Johnson 2000). On this view, the argument-product, a package of reasons supporting a conclusion (presented in the form of a text, for example), issues from a social *process of arguing* that is more or less regulated by formal or informal *procedures*—for example, rules that define turn-taking, burdens of proof, admissible statements, and so on. Such procedures define who may say what to whom, and when they may say it, in the course of an argumentation process.[8]

Within this framework, one can define the cogency of argument-products in different ways.[9] On the one hand, the theorist can ask whether cogency depends on features of the product alone, or also on the quality of the procedure, or still further on contingent social-psychological and historical features of the audience, that is the broader social process as it eludes formal proceduralization. On the other hand, the particular normative standards that are applied to each of these dimensions can vary according to one's analytic perspective. As described above, the logical empiricists focus on the properties of argument-products, whose cogency they define solely in light of formal-logical standards, whereas Achinstein regards both logical and representational properties as essential.

However, both approaches also suggest that practical relevance requires one to go beyond the focus on argument-products to include processes of argumentation. Achinstein's model implies that the assessment of an argument can only lead to a defeasible estimate of cogency. The information needed for the assessment of an argument will not always be included in the

---

[8] See Wenzel (1990). Insofar as the product is an identifiable text, one can distinguish product from process rather straightforwardly. The procedure-process distinction is more difficult, given that procedures are embedded in complex social processes; "process," then, is the more encompassing of the two terms, and we use the term "procedure" only to single out more explicitly rule-guided aspects of process.

[9] In speaking of "defining cogency," we have a broadly normative sense of definition in mind that would encompass any project—including those at issue here: logical explication in Hempel and Carnap's sense, Achinstein's metaphysical definition, and pragmatic attempts to articulate norms of argument-making practices--that issues in a statement of the properties in virtue of which we ought to regard an argument as cogent.

argument-product itself. On Carnap's model also, such an estimate, even when calculated correctly, faces the possibility of being overthrown should it turn out that the total evidence requirement is not after all satisfied. But bringing background assumptions to the fore for open criticism typically requires an ongoing intersubjective process of critical testing in which arguments confront challenges of various sorts. Thus Achinstein's "potential empirical incompleteness" (and to an extent Carnap's total evidence requirement) requires us to move beyond a narrow product-centered, inductive logic-based model of cogency.

The practical, intersubjective contexts of argumentation have garnered considerable attention from argumentation theorists. Within science studies, the rhetorical perspective has probably received the most attention. But as we shall see, the dialectical perspective plays an important role at CDF. In argumentation studies, the dialectical tradition has been associated with formal debate procedures and, more broadly, with processes of critical discussion in which participants test claims and arguments for their ability to hold up to challenges.[10] More recently, argumentation theorists have linked dialectic with normative dialogical or conversational models of argumentation. These models are often proposed as a kind of "informal" logic (Walton 1989; 1998; Eemeren *et al*. 1993).

Argumentation studies thus allows us to reformulate our leading question more precisely, as a theoretical challenge that confronts objective theories of evidence: how, exactly, are the practical social contexts of evidential arguments related to the cogency of those arguments? If we distinguish between the *definition* of cogency and its *assessment*, then two possible answers

---

[10] See Rescher (1977); Popper (1989, 313 note 4) associates the trial and error method with classical dialectic in the sense of argumentative testing; for the roots in Aristotle and an application to Galileo's argumentation, see Zuber (1998); for an attempt to develop a dialectical model of science, see Pera (1994).

immediately present themselves. The adapted theories of evidence described above appear to define cogency entirely in terms of impersonal merits, as distinct from its practical assessment. However, work on argumentation suggests that one might, alternatively, define cogency as at least partly dependent on facts about persons involved in the construction or reception of the argument. For each camp, the question of assessment arises – an assessment that may well require consideration of intersubjective processes of argumentation, either as constitutive of the cogency of the argument at hand (the second answer above) or as extrinsic indicators of the impersonal merits of the argument (the first answer). Thus on either hand one must confront the question of what kinds of argumentation processes are relevant to the assessment of cogency. Do we invoke dialectical, rhetorical, or social-institutional perspectives—or a combination thereof?

We address these questions in the next two sections, with the help of an actual case. This case will help to illustrate some subtle difficulties facing the attempt to define cogency in impersonal terms, while also showing how an argumentation-theoretic approach can help illuminate the social practices of scientific argumentation. Of particular relevance will be (1) the use of *probation by process*, in which knowledge of the dialectical process by which an argument was constructed plays a role in the assessment of the argument's cogency, and (2) the achievement of *heterogeneous consensus*, in which a collaboration agrees to the publication of an evidence claim, while disagreeing on the premises offered in that publication as support for the claim.

2. The CDF Collaboration: Procedures that Structured the Writing Process

Our case concerns an important evidential argument advanced by the Collider Detector at Fermilab (CDF) collaboration in a paper announced to the public in April 1994 (Abe et al. 1994). This paper (henceforth "the Evidence paper") put forth the first claim to have found evidence for the existence of the top quark, the last of the six "flavors" of quarks included among the elementary particles in physicists' Standard Model to be experimentally confirmed.[11] After a brief introduction to the physics background (2.1) we describe the semi-formal procedures employed at CDF (2.2) and then assess the impact of these on cogency (2.3, 2.4).

(2.1) In 1994, the CDF collaboration had a membership of about 450 physicists from 34 institutions representing five different countries. Their massive and complex detector sat at a point where beams of protons and antiprotons, brought to very high energies by Fermilab's Tevatron accelerator, collided head-on. The barrel-shaped detector surrounded the collision point and was made up of many different kinds of components designed to measure different properties of particles resulting from the proton-antiproton collisions.

If the top quark existed, then CDF expected it to be produced in top–anti-top pairs in those collisions at some rather low rate negatively correlated with the unknown rest mass of the top quark. The top would almost immediately decay into other particles, and the experimental search would proceed by looking for the characteristic "signature" decay products of the top quark. CDF rendered that signature precise by establishing "cuts" on the specific particle measurements made by the detector, although those cuts could still be satisfied by other "background" particles. In the "counting experiment" approach, the experimental problem was thus translated into a partly statistical one – how to choose a sensible set of cuts, such that by counting the number of events passing the cuts ("candidate events"), one could reliably evaluate whether the number of candidate events in a given data set significantly exceeded the number expected from background to an extent that could be taken as an indication that the top quark was contributing to the decays registered in the detector.

[11] The discussion here draws from the much more detailed discussion in Staley 2004a, in which the evidential issues are discussed within the context of the error statistical theory of evidence (Mayo 1996).

CDF developed three different counting experiment analyses that looked for two different decay modes of the top quark. The "dilepton" counting experiment looked for events with two leptons with a large "transverse momentum" (the component of the particle's momentum perpendicular to the beam line), along with two or more energetic "jets" of strongly interacting hadrons. The other two counting experiments both looked for events with one high transverse momentum lepton, three or more energetic momentum jets, and some indication that the event contained a bottom, or *b,* quark. The Secondary Vertex Tagging experiment (SVX) used the presence of a jet originating from a point removed from the main event collision ("primary vertex") to tag events in which a *b* quark had been produced close to the primary vertex, traveled a short distance, and then decayed, resulting in a jet. The Soft Lepton Tagging experiment (SLT) relied on an identification of a low transverse-momentum ("soft") lepton as a way to tag events in which the top quark had decayed to a *b* quark and a massive W boson, and the *b* quark further decayed to produce a lepton.

(2.2). The top quark was not CDF's only experimental concern, and by the time of the announcement of the Evidence paper, CDF had well-established routines for the production of experimental results and the writing of papers reporting those results.

Within CDF were numerous "physics working groups" devoted to particular kinds of physics problems and analysis. The working groups were further divided into subgroups that were successively more specialized in focus. Thus the "heavy flavors" group was concerned with physics problems relating to the bottom and top quarks. A large top quark working group constituted part of the heavy flavors group, and the top group itself was further divided into smaller subgroups that focused on particular strategies for identifying top quark decays, such as the "dilepton group."

CDF's published papers list most collaboration members as authors, but the initiating steps that lead to publication begin within the much smaller working groups, in which physicists actually develop the analyses that potentially lead to publishable results. When an analysis is sufficiently far along to contemplate a publication, the leaders of the working group approach the collaboration spokespersons (there are two and they are elected) to request the appointment of a "godparent" committee.

The godparent committees are meant to function as internal referees who are responsible for careful scrutiny and criticism of any work that is to be proposed as a publishable result by the collaboration. Only after the godparents give their approval can the working group present a draft of a paper to the collaboration as a whole. At that time, the paper is posted electronically to the collaboration, and all members of the collaboration are free to ask questions and raise criticisms of the paper. The originators of the analysis are given the opportunity to respond to such comments and revise their paper as needed. The decision to publish is governed by a rule of consensus, but perfect unanimity is not required. Individual dissidents may choose to remove their name from a paper, although this is perceived as a measure of last resort.[12]

Notice that this institutional framework, in which the Evidence paper was developed, provided a set of dialogical procedures for regulating the collective intellectual process of producing a publishable argument. As a collective effort to generate an argument—an intellectual product—both social and intellectual factors intertwined in a *dialectical* process of critical testing. Accordingly, the procedures we have just described parceled out the intellectual aspects of the writing process (first draft production, initial scrutiny, open critical assessment) across different actors (the working group, spokespersons, godparents, collaboration as a whole).

---

[12] CDF members may also decline to list themselves as authors on a paper when they feel that they did not make a significant contribution to the work presented, or are not sufficiently familiar with the contents of the paper. The decision is left to the individual physicist, standards vary, and the issue is a matter of some controversy among collaboration members.

However, in the case of the Evidence paper, some details of its history—how the procedures were actually conducted—are peculiar, and the paper as a whole was significantly more controversial, more publicized, and longer than most CDF publications. After the appearance of excesses beyond background expectations in all three counting experiments in July 1993, the members of the top quark working group asked for and received a godparent committee for the top quark results. This was unusually early, in that they had not yet prepared a draft of any paper.

The collaboration initially clashed over the choice of a format for presenting the results. One possibility was to publish their results in four short papers in *Physical Review Letters*, one for each of the three counting experiments, with the fourth paper giving a brief summary of the combined counting experiment results as well as an estimate of the top quark mass. Instead, the collaboration directed the top group to write a single, detailed paper for publication in *Physical Review D*. Collaboration members Tony Liss and Paul Tipton led a core group of about ten people who met a couple of times a week from the time of the October collaboration meeting until the release of their first draft. Each member institution then jointly posted their criticisms and questions in a computer file open to the whole collaboration. Liss and Tipton, who were responsible for posting replies (though others in the core group wrote many of the replies) took the unusual step of sending their replies first to the godparents, who reviewed them before sending them on to the collaboration as a whole, sometimes intervening to impose a more courteous tone than had been achieved by the weary and overburdened core group.

(2.3) The Evidence paper drew an unusual amount of attention within the collaboration from members who did not have a direct hand in developing the analysis. Moreover, because the paper drew together several different analyses that had been developed by different subgroups, those who had played some significant role in the analysis comprised a diverse group with widely varying perspectives. By April 1994, when the collaboration had reached a consensus that they were ready to publish, many collaboration members, including some who had played central roles in the top quark analysis from the beginning, found that their views of the evidence had changed over time, and had done so in part as a result of the dialectical structure of the writing procedure. Noted one prominent collaboration member, "even if I hadn't been paying attention to all of the detail . . . I would have been more confident [of the evidence claim], just having sat in all of the meetings, in which so many people asked so many questions, demanding that they be answered at a quantitative level. . . . I was much more confident at the end" (Staley 2004a, 169).

We regard this strengthening of confidence as an outcome of the collaboration's institutional procedures as an instance of *probation by process*, and in this case it is no accident. The very purpose of the godparent committees is to enable the production of stronger evidential arguments that will meet the consensus standard within the collaboration for publication, as well as generate acceptance in the broader community after publication. (Some collaboration members expressed the view that the criticism from inside the collaboration is much more stringent and well-informed than anything encountered "on the outside.") In section three we will consider the implications of probation by process for the idea of cogency.

(2.4) The story told thus far about the writing of the Evidence paper by CDF has glossed over a large number of complications and plot twists. Without delving into the details, it will be worth seeing how controversy over the status of different aspects of the experiment arose, and how (and to what extent) that controversy was resolved in order to reach agreement on a publication.

In addition to the counting experiments described above, the top group included at least three other groups working on different versions of a "kinematic" analysis approach to finding the top quark. These analyses tended to make use of a likelihood comparison – i.e., a comparison of the probabilities of the kinematic characteristics of selected events under the hypotheses of the event being due to top decay and of the event being background. Such a comparison could then be used to define some measure of "top-likeness" for the event in question.

During the production of the Evidence paper, a dispute emerged over the role of the kinematic analyses in the Evidence paper that, in the words of one former spokesperson, "almost split the organization open" (Staley 2004a, 150). The route eventually chosen was a kind of compromise in which a section of the paper discusses the kinematic features of the events chosen by the counting experiments, and compares those features with predictions from simulations of top quark decays. However, this information is not part of the central result claimed in the paper, and is not used in evaluating the statistical strength of the evidence. The kinematic features of the events are treated as a "cross-check" on the central evidence claim.

Collaboration members, in subsequent interviews, continued to diverge considerably in their estimate of the evidential import of these kinematic features of the events. Some felt that the results of the kinematic analysis should have been presented alongside and statistically combined with the counting experiment results. Others thought it was appropriate to include as either supporting evidence or as a cross-check, but not as part of the central evidence claim. Still others believed that it had been included only as a political expedient to keep the writing process moving forward. They did not trust the kinematic analyses to deliver reliable results, and were troubled by the degree to which these analyses depended on complicated "Monte Carlo" computer simulations of background processes.

On the other hand, the counting experiments had their own problems. Worries arose for some group members over possible biases in both the SLT and SVX searches, though those worries had very different sources in the two cases.

The SLT analysis had been carried out by a very small number of physicists working in relative isolation from other group members. Well into the analysis of data from the 1992–93 run, the value of the minimum threshold for the transverse momentum of the soft lepton had been left undecided between two values. The fact that the number of SLT candidate events depended strongly on which value was chosen (seven events versus four) raised concerns that the published SLT cuts, which yielded the larger number of candidate events, were chosen *because* they resulted in a larger number of candidate events.

The SVX analysis had emerged from a much larger subgroup that had pursued multiple algorithms. In fact, three different SVX algorithms had been developed by three different "sects" (as one group member called them). As in the case of the SLT analysis, this development process overlapped considerably with the actual analysis of already collected data. Some collaboration members formed the perception that the SVX group had taken on the dynamics of a horse race, with the winning group to be determined by whichever algorithm selected the largest number of candidate events.

At the time, none of these worries over bias could be substantiated, and both the SLT and SVX groups could respond with prima facie plausible – but also not completely conclusive – defenses against the charges that their counting experiments were biased.[13]

Given these disagreements over the different aspects of the Evidence paper, how could anyone say that CDF had produced a consensus? In subsequent interviews, when CDF physicists were asked whether they still endorsed the claims made in the Evidence paper, they typically replied that they did, and then many went on to explain how one or another part of the analysis was flawed, but that other aspects of the analysis had in fact been evaluated conservatively (or not included in the central evidence claim at all), so as to compensate for those flaws. Different critics, however, identified different strengths and weaknesses. Thus, although the collaboration

[13] A later data run yielded results that were at least consistent with there being some biases in the earlier counting experiments – though this also was not conclusive and yielded no information at all about the source of any biases. The issues are in fact quite complex; see Staley 2004a, esp. ch. 5.

reached a general consensus on their evidence claim, this was a *heterogeneous consensus* insofar as the reasoning on the basis of which this claim was accepted varied from one individual to another.

The fact that most experimental work is done in collaborations is a point that should not be overlooked by argumentation theorists seeking to make sense of argumentation practices in the sciences. This fact means that the argument-products that one finds published in scientific journals must typically be understood as products of a political process. What does the variety of views within CDF about the Evidence paper argument mean for the evaluation of that argument's cogency?

3. Discussion: From Impersonal to Dialogical Merits

Here we consider two features of the story of CDF's Evidence paper – their use of probation by process and their reliance on a heterogeneous consensus – and put them into argumentation-theoretic perspective. In so doing, we discuss their implications for the concept of argumentative cogency.

(3.1) To understand the use of probation by process in constituting or assessing cogency, we will need to distinguish a generative perspective from an evaluative one, with the role of procedure figuring differently in the two cases. From a generative perspective, procedure plays an *instrumental* role with respect to cogency. By requiring their papers to go through instituted mechanisms of criticism and dialogue, CDF reduces their risk of publishing papers that turn out to be less cogent than they believe them to be. If cogency is defined in terms of the impersonal merits of the argument product, then we might assume that dialectical testing procedures helped (a) to eliminate flaws in the logical structure of the paper's argument, (b) to identify those evidential reasons that provided the best support for the claim to have evidence for the existence of the top quark, that is, the reasons that were most plausible (probably true, empirically adequate) and most likely to count as evidence in the light of background considerations. The argument's cogency is a matter of its providing evidence that a sufficiently wide variety of possible experimental and logical flaws have been ruled out. Whether a particular argument has met this standard is understood on this view as a fact that obtains without respect to the authors, the audience, or any other persons. For collaboration members seeking to construct an argument that can withstand critical scrutiny, it is advantageous to have in place a set of dialectical procedures that minimize the chance of generating arguments that fail to meet this standard. The value of such procedures of severe criticism is enhanced by the fact that whether the standard is met cannot in general be read directly off of the text of the argument.

From the perspective of a person considering an argument and trying to evaluate its cogency—the audience member as spectator-judge—the cogency-as-impersonal-merits view suggests a different role: Procedures serve as a source of indirect *evidence* with respect to the cogency of the arguments they produce. Although the cogency of the argument is, again, a matter of the satisfaction of a number of background assumptions and the absence of any evidence-defeating interference, the knowledge that the argument-product at hand passed through rigorous dialectical procedures of argument construction and refinement serves as an additional although indirect evidential resource. The evaluator can reason that, were there important countervailing considerations, they would have been brought to light before the argument-product was allowed to be presented in its current form, a fact that again cannot be ascertained directly from the text of the argument itself. The spectator-evaluator, in other words, employs probation by process in an *indirect* assessment of the argument's impersonal merits.

On the other hand, if one defines cogency in such a way as to allow that merits of an argument may involve facts regarding persons, then from both the generative and evaluative perspective, CDF's institutional procedures partly *constitute* the cogency of the argument. We might then regard the cogency of an argument as determined in part by *dialogical* merits, i.e., features of the argument that enable it to survive critical questioning by a given audience. The idea is that the product of argumentation reflects the quality of the process, such as the level of the critical testing, the social-psychological conditions that foster or hinder responsible judgments on the part of the participants, and so on. Some of these merits are audience- or discipline-specific conventions of representation and the like, others are features that one can in principle analyze impersonally, but whose cogency depends on the audience.[14]

It is worth noting, as a possible consideration in support of such a unification of the generative and evaluative perspectives, that the distinction between them is not easily drawn in the context of a large collaboration producing a complex argument such as the evidential argument for the top quark. Work was distributed piecemeal amongst the collaborators. No collaboration member was in the position of being able to think of procedures underlying the textual materials at hand (including those constituting "raw materials" for the article in preparation, such as plots and data summaries) in purely instrumental terms. Rather, because each *part* of the argument drew upon the combined efforts of many individuals, the procedures employed were potentially relevant to the evaluation of claims made in the text even for those collaborators who were primarily responsible for the presence of those claims in the text. In effect, every collaborator, including those actively involved in writing the text, was in the audience-spectator position in assessing the entire paper.

(3.2) According to some models of argumentative practices (e.g., Habermas 1990), dialogical success in the full sense means that the interlocutors accept the conclusion on the basis of the same reasons, and they agree on the force of that conclusion, that is, that it is beyond reasonable doubt, say, or simply has enough support to merit further study. Although some

[14] E.g., the level of statistical significance for publishable findings and the depth of explanation of specific methodologies, vary with audience and discipline.

theorists of science (e.g., Solomon 2001) have questioned consensus as a goal of inquiry, the CDF procedures and debate aimed at a consensus of the sort Habermas proposes, that is, they wanted to reach agreement on both the conclusion and the evidence that provided reasons supporting the conclusion.[15] In reaching only heterogeneous consensus, the CDF Collaboration fell short of reaching this goal for the Evidence paper (although their next paper presenting data

[15] The goal of consensus at CDF is tied up with the demands of responsible authorship: each physicist, in claiming authorship, took responsibility for the content of the paper; Solomon, by contrast, is concerned with the fruitful distribution of research agenda across an entire discipline of field.

on the top quark reflected a more uniform agreement; see Staley 2004a, chap. 5).[16] What difference does it make to our account of heterogeneous consensus whether cogency is defined in impersonal or dialogical terms?

If one defines cogency in terms of the impersonal merits of the product, and considers procedure and process as only indirectly relevant for assessing such merits, then a heterogeneous consensus could signal one (or both) of two sources of failure: (a) that the proffered arguments lacked sufficiently strong impersonal merits to garner a thick, or "homogeneous" consensus; (b) the process of presentation and discussion behind the arguments was poorly designed or executed (rhetorical misfires in the text, dubious behavior calling into question the trustworthiness or good

[16] One might be tempted to fault CDF on this point, as it means that, in effect, they did not apply the same standard to their premises as they did to their conclusion. Two points are worth noting in CDF's defense, however: (1) CDF physicists who rejected claims relating to some part of the analysis that was officially included as part of the evidence for the top quark defended their support for the paper on the grounds that other parts of the analysis yielded compensating evidence that was in fact stronger than the paper claimed. (2) A growing literature on the "discursive dilemma" indicates that groups seeking to aggregate individual judgments on a set of logically related propositions must, to put it roughly, *choose* between (a) using a uniform method to aggregate individual judgments on each proposition, and (b) ensuring that the judgments reached in the aggregate meet the minimal rationality standard of logical consistency (see, e.g., Pettit 2001, List and Pettit 2002, Pauly and van Hees 2006, List 2006). From a positive perspective, one might take the possibility of sustaining the result on the basis of different supporting reasons as an indication of the *robustness* of that result, in the sense discussed by Wimsatt 1987 and Staley 2004b.

faith of participants[17], etc.). If the first possibility explains the outcome, then the heterogeneous character of the consensus is a symptom of a lack of cogency of the arguments. In the case of the top quark Evidence paper, for example, one could infer from the heterogeneous consensus that the arguments, taken together, were sufficiently strong to engender consensus on a provisional conclusion, namely that evidence for the top quark exists but further research is required to establish the top as a fact. But the impersonal merits were not strong enough to garner consensus on the evidence itself, that is, consensus on which reasons were in fact the right reasons. Note, however, that to make this indirect assessment, one must presuppose that the process and procedures of discussion were sufficiently meritorious.

If, on the other hand, problems of process explain the failure to reach a homogeneous consensus on the argument (alternative b above), then the assessment of the argument in terms of impersonal merits becomes more problematic, requiring one, in effect, to consider the specific difficulties of process that were encountered and consider whether, had things gone better, the impersonal merits of the argument would have been sufficiently strong to produce a homogeneous consensus. In either case, although on this view process considerations do not constitute the cogency of scientific arguments, the evaluation of such arguments under conditions of heterogeneous consensus calls for some sort of assumption about the processes that contributed to the construction of the argument.

On the second view, the relevance of the dialogical merits of the argument generation

[17] Commenting on her perception that some collaborators behaved in a manner suggesting a degree of possessiveness regarding aspects of the analysis in the top quark evidence paper, one CDF physicist commented, "Unfortunately, that whole era was blurred by what I considered pretty bad behavior of people. . . . You would ask questions and a lot of times the answers were tinged with a kind of anger. . . . That just makes you want to believe it less" (Staley 2004a, 170).

process is built into the understanding of cogency. Of course, the same possibilities obtain with regard to the explanation of heterogeneous consensus – it could be the product of flaws with regard to impersonal merits, dialogical merits, or both.

We thus have two theoretical approaches to argumentative cogency, which we might label the *objectivist* and the *dialogical* models. The former does not deny the relevance of dialogical merits, but regards their relevance as indirect and extrinsic to cogency itself. The latter does not deny that arguments have impersonal merits or that scientists seek to establish the truth of their hypotheses. Rather, the model operationalizes the search for truth in terms of a search for increasingly cogent arguments, where cogency includes dialogical merits. Thus this brand of operationalism is not so much anti-realist or social constructionist as pragmatist in spirit.

The objectivist and the dialogical models appear to be operationally equivalent in the sense that a person attempting to assess the cogency of a given argument will under either theory need to consider the very same kinds of facts, including both impersonal and dialogical factors. Where the two theories differ is with regard to whether the dialogical factors relate intrinsically or extrinsically to the content of the judgment of cogency that is reached. Under both accounts one must make judgments about the processes through which the argument in question was built. Procedural helps and process dynamics are thus indispensable for real human knowers.

(3.3) To make good on our aim of practical relevance for scientists, we offer a practical proposal. Our proposal is merely preliminary and somewhat broad, but we regard it as grounded in reflection on the role of process considerations in the case of the CDF Evidence paper.

Many, though certainly not all, of the controversies that surrounded the claims put forth in the Evidence paper turned on questions about the decision procedures used to determine data selection criteria – procedures that may have introduced biases into the statistical analysis of the results. Skepticism lingered where answers to such questions were either not convincing or not reassuring. Such problems were not unique to CDF. The CDF collaboration shares an accelerator facility with another large colliding beam detector collaboration, D-zero. Physicists interviewed in each collaboration expressed skepticism about some results published by the other group, and these doubts often arose on the basis of a perception that decision procedures for setting data selection criteria were obscure and potentially biased.

Of course, physicists generally, and high energy experimental physicists in particular, are aware of this issue, and are generally aware of the organizational issues implicated. One result is a growing use of “blind analysis” techniques that in various ways prevent decisions about selection criteria from being influenced by the data to which those criteria will be applied (Franklin 2002). Such techniques, however, are not always applicable or practical, and so we propose that a more fundamental and general response should be for experimental collaborations to adopt policies that will make their decision-making procedures more “transparent” to their colleagues seeking to understand and evaluate specific experimental arguments. The details of how such a reform might be carried out will be difficult to negotiate, as such groups, although they may be eager to defend the cogency of their arguments, are often also eager to protect their exclusive access to their data and to keep “internal” disputes and difficulties, which often have a strong personal component, from spilling out into the public domain. Here we simply note that, however such tensions might be resolved, these groups already keep rather copious records for their own purposes, and often have their own computer network accessible only to their own members on which such information is kept. Improving the ability of the community to become informed of the process considerations relevant to particular experimental arguments can thus be achieved by enhancement and modification of practices already in place.

Of course we are not here proposing something that no experimentalist has ever considered. But CDF physicists interviewed about these issues often classified them as concerning “sociology” rather than physics. To return to our philosophical question regarding the place of process considerations in the definition of argumentative cogency, perhaps the strongest consideration in favor of including dialogical merits in the definition of cogency is that such a view makes the implementation of cogency-enhancing procedures *internal* to scientific activity. That is, if adopted, such a view makes good procedures simply *part of* good scientific argumentation, rather than a merely instrumentally helpful adjunct to science proper. The latter view, while certainly not incoherent, has the potential practical weakness of making it easier for scientists to neglect process considerations, resulting in greater controversies and disputes over particular results.

Such an approach also fits well with empirical work by Kevin Dunbar and collaborators on collective scientific reasoning. Dunbar, using a combination of traditional cognitive psychology experimentation ("in vitro") and direct observation of small scientific collaborations engaged in problem solving ("in vivo"), presents a naturalistic view in which distributed reasoning figures as a "key cognitive feature" of the category science alongside such reasoning activities as analogy, deduction, and induction. Like the view we are suggesting, Dunbar's approach thus gives the process of collaborative reasoning a cognitive (and in our terms cogency-relevant) status comparable to that of other widely-recognized inferential activities.

(3.4) To conclude, we notice that these alternative models of cogency, namely the objectivist and dialogical, require the theorist to make a philosophical decision over how to regard the concept of argumentative cogency: objectivism or pragmatism?

Both sides acknowledge points made by the other but situate them differently in relation to the definition of cogency. The objectivist grants the importance of dialogical process: not only formal procedures designed to enhance the dialectical strength of arguments, but also social-psychological considerations that affect scientists' ability to decide whether the appearance of cogency might be illusory or not. But objectivists insist that a distinction be drawn between cogency as such and process-based sources of evidence used to assess the cogency of particular arguments—the role played by process remains indicative, not constitutive.

Pragmatists can grant the objectivist point that truth, or fidelity to facts of nature, constitute merits of scientific argument. But they deny that there is a fact of the matter whether any argument-product, taken on its own apart from process, is cogent. The pragmatist strategy thus declines to draw the distinction that is central to the objectivist interpretation. Since cogency cannot be evaluated without at least asking questions about possibly relevant process considerations, one can dispense with the abstraction required by the objectivist and simply define cogency to include those aspects of inquiry that inspection of the argument does not disclose.

In calling the decision between the objectivist and pragmatist strategies philosophical, we do not mean to say that it is merely academic. Quite to the contrary, decisions about theory-construction can be made for better or worse. A complete weighing of likely consequences of these alternative strategies in theorizing about argumentative cogency is beyond the scope of this paper. We simply noted that one possible consequence regards the way scientists view process considerations: either as merely "sociological" or as constitutive of good scientific argumentation. The desirability of the latter view is itself a pragmatic consideration in favor of the pragmatic over the objectivist approach to cogency. Whether there are additional epistemic (truth-related) reasons in favor of the pragmatic approach to cogency we leave as an open question. What an argumentation-theoretic approach to the CDF collaboration shows is that, on either approach, a great deal of attention will need to be paid to the contributions of social process in the weighing of argumentative cogency. Both the objectivist and the pragmatist must still draw upon the same argumentative phenomena when asked to measure the merits of argument.

Reference List

Abe, F.,, et al. [CDF Collaboration]. 1994. "Evidence for Top Quark Production in *p-pbar* Collisions at √*s* = 1.8 TeV." *Physical Review D* 50: 2966–3026.

Achinstein, Peter. 1991. *Particles and Waves*. New York: Oxford.

———. 2000. "Why Philosophical Theories of Evidence Are (and Ought To Be) Ignored by Scientists." *Philosophy of Science* 67 (Supplement: Proceedings oof the 1998 Biennial Meeting of the Philosophy of Science Association):S180–92.

———. 2001. *The Book of Evidence*. New York: Oxford University Press.

Bloor, David. 1991. *Knowledge and Social Imagery,* 2nd ed. Chicago: University of Chicago Press.

Carnap, Rudolf. 1962. *Logical Foundations of Probability*. Chicago: University of Chicago Press.

Dunbar, Kevin. 2002. "Understanding the Role of Cognition in Science: The *Science as Category* Framework." In Pp. 154–70 in *The Cognitive Basis of Science*. Edited by Peter Carruthers, Stephen Stich, and Michael Siegal. Cambridge: Cambridge University Press.

Eemeren, Frans H. van, Rob Grootendorst, Sally Jackson, and Scott Jacobs. 1993. *Reconstructing Argumentative Discourse*. Tuscaloosa: University of Alabama Press.

Eemeren, Frans H. van, Rob Grootendorst, Francisca Snoeck Henkemans, J. Anthony Blair, Ralph H. Johnson, Erik C. W. Krabbe, Christian Plantin, Douglas N. Walton, Charles A. Willard, John Woods, and David Zarefsky. 1996. *Fundamentals of Argumentation Theory*. Mahwah, NJ: Erlbaum.

Franklin, Allan. 2002. *Selectivity and Discord: Two Problems of Experiment*. Pittsburgh: University of Pittsburgh Press.

Govier, Trudy. 2005. *A Practical Study of Argument*. 6th ed. Wadsworth.

Gross, Alan G. 2006. *Starring the Text: The Place of Rhetoric in Science Studies*. Carbondale:

Southern Illinois University Press.

Habermas, Jürgen. 1990. "Discourse Ethics: Notes on a Program of Philosophical Justification." Pp.43–115 in *Moral Consciousness and Communicative Action*. Trans. C. Lenhardt and S. W. Nicholsen. Cambridge, Mass.: MIT Press.

Harding, Sandra. 1998. *Is Science Multicultural?* Bloomington: Indiana University Press.

Harris, Randy Allen. 1997. *Landmark Essays on Rhetoric of Science: Case Studies*. Mahwah, NJ: Hermagoras.

Hempel, Carl G. 2000. "The Problem of Truth." Chap. 4 in *Selected Philosophical Essays*. Edited by R. Jeffries. Cambridge: Cambridge University Press. First published in 1937.

Hempel, Carl G. and Paul Oppenheim. 1945. "A Definition of 'Degree of Confirmation'." *Philosophy of Science* 12:98–115.

Johnson, Ralph H. 2000. *Manifest Rationality: Toward a Pragmatic Theory of Argument*. Mahwah, NJ: Erlbaum.

Kitcher, Philip. 1993. *The Advancement of Science*. New York: Oxford University Press.

———. 2001. *Science, Truth, and Democracy*. Oxford: Oxford University Press.

Klein, Julie Thompson. 1996. *Crossing Boundaries: Knowledge, Disciplinarities, and Interdisciplinarities*. Charlottesville: University Press of Virginia.

Kusch, Martin. 2002. *Knowledge by Agreement*. Oxford: Clarendon.

List, Christian and Philip Pettit. 2002. "Aggregating Sets of Judgments: An Impossibility Result," *Economics and Philosophy* 18:89-110.

List, Christian. 2006. "The Discursive Dilemma and Public Reason." *Ethics* 116:362–402.

Longino, Helen E. 1990. *Science as Social Knowledge.* Princeton: Princeton UP.

———. 2002. *The Fate of Knowledge.* Princeton: Princeton UP.

Mayo, Deborah. 1996. *Error and the Growth of Experimental Knowledge*. Chicago: University of Chicago Press.

Pauly, Marc and Martin van Hees. 2006. "Logical Constraints on Judgment Aggregation," *Journal of Philosophical Logic* 35:569–85.

Pera, Marcello. 1994. *The Discourses of Science,* trans. C. Botsford. Chicago: University of Chicago Press.

Pettit, P. 2001. "Deliberative Democracy and the Discursive Dilemma," *Philosophical Issues* 11:268–99.

Pickering, Andrew, ed. 1992. *Science as Practice and Culture.* Chicago: University of Chicago Press.

Polanyi, Michael. 1951. *The Logic of Liberty.* Chicago: University of Chicago Press.

Popper, Karl R. 1957. *The Poverty of Historicism.* London: Routledge.

———. 1989. *Conjectures and Refutations.* 5th ed. London: Routledge.

Rehg, William. 1997. "Reason and Rhetoric in Habermas's Theory of Argumentation." Pp. 358–77 in *Rhetoric and Hermeneutics in Our Time.* Edited by W. Jost and M. J. Hyde. New Haven: Yale University Press.

Rescher, Nicolas. 1976. *Plausible Reasoning.* Assen/Amsterdam: Van Gorcum.

———. 1977. *Dialectics.* Albany: SUNY Press.

Salmon, Wesley. 1967. *The Foundations of Scientific Inference*. Pittsburgh: University of Pittsburgh Press.

Solomon, Miriam. 2001. *Social Empiricism*. Cambridge: MIT Press.

Staley, Kent. 2004a. *The Evidence for the Top Quark: Objectivity and Bias in Collaborative Experimentation*. New York: Cambridge University Press.

———. 2004b. "Robust Evidence and Secure Evidence Claims." *Philosophy of Science* 71:467–88.

———. 2005. Agency and Objectivity in the Search for the Top Quark." Pp. 165–83 in *Scientific Evidence: Philosophical Theories and Applications*. Edited by Peter Achinstein. Baltimore: Johns Hopkins University Press.

Walton, Douglas. 1989. *Informal Logic*. Cambridge: Cambridge University Press.

———. 1998. *The New Dialectic*. Toronto: University of Toronto Press.

Wenzel, Joseph. 1990. "Three Perspectives on Argument." Pp. 9–26 in *Perspectives on Argumentation: Essays in Honor of Wayne Brockriede*. Edited by R. Trapp and J. Schuetz. Prospect Heights, IL: Waveland.

Wimsatt, William. 1987. "Robustness, Reliability, and Overdetermination." Pp. 124–63 in *Scientific Inquiry and the Social Sciences*. Edited by M. Brewer and B. Collins. San Francisco, Jossey-Bass.

Zuber, Marta Spranzi. 1998. "Dialectic, dialogue, and controversy: The case of Galileo." *Science in Context* 11:181-203.