volume II ♦ issue 3

2011

Logos & Episteme

an international journal
of epistemology

Romanian Academy
Iasi Branch
◆◆◆
"Gheorghe Zane" Institute
for Economic and Social
Research

INSTITUTUL EUROPEAN

Contact address:

Institutul de Cercetări
Economice şi Sociale „Gh.Zane”

Iaşi, str.T.Codrescu, nr.2, cod 700481
Tel/Fax: 004 0332 106507
Email: logosandepisteme@yahoo.com

www.logos-and-episteme.proiectsbc.ro

Logos & Episteme

ISSN-L: 2069-0533
ISSN: 2069-0533

This journal was edited within The Knowledge Based Society Project supported by the Sectoral Operational Programme Human Resources Development (SOP HRD), financed from the European Social Fund and by the Romanian Government under the contract number POSDRU ID 56815

TABLE OF CONTENTS

Articles

PAUL RICOEUR'S HERMENEUTICS BETWEEN EPISTEMOLOGY AND ONTOLOGY*

Cătălin BOBB

ABSTRACT: The aim of our text is to explore the ties of Ricoeur's hermeneutics with ontology and epistemology. We have to admit that (1) for Ricoeur, at the beginning of his work, hermeneutics (as one can find it in *Le conflit des interprétations*) was never a main topic (hermeneutics, as hermeneutic intelligence, was always a solution to a certain problem and never a problem in itself), and (2) that when hermeneutics becomes a main topic (as one can find it in *Du texte à l'action*), the purpose of Ricoeur is to suggest a renew ontological hermeneutics, beyond Heidegger and Gadamer, but still tied with his non-hermeneutic intents. Our thesis is that Ricoeur's latest hermeneutics, beyond his epistemological status, can be regarded as ontology. Of course, one cannot find a direct ontology, as we can find it in Heidegger or Gadamer, but one can find what we can call a reversed ontology, an ontology which does not start from the centre of the human experience of understanding but outside of it. In other words, we are going to show that not even his late hermeneutics (the critical moment), better known as textual hermeneutics, is not per se an epistemological hermeneutics beyond its declared intention as being one.

KEYWORDS: Paul Ricoeur, hermeneutics, epistemology, ontology

1. Introduction

For methodological reasons, we have to state what we understand by *hermeneutics*. From the vast literature on this topic we can adopt two main definitions of hermeneutics: hermeneutics as the technique of interpretation (a clear set of rules) applied on texts, facts, history, and so on, and second, hermeneutics as the problem of understanding, or, under the account of Jean Greisch, the problem of understanding the understanding.[1] Of course, the second characteristic of

* This paper was made within *The Knowledge Based Society Project* supported by the Sectorial Operational Program Human Resources Development (SOP HRD), Financed by the European Social Fund and by the Romanian Government under the contract POSDRU/89/1.5/S/56815.

[1] Jean Greisch, *Herméneutique et grammatologie* (Paris: CNRS, 1977), 25.

hermeneutics responds to what Jean Grondin calls *philosophical hermeneutics* which, in his view, begins with Gadamer and, in some aspects, with Ricoeur.[2]

Even so, if we are going to view Ricoeur's hermeneutics between these two features, some difficulties may occur. I want to emphasize only one key difficulty. The two main volumes that constitute the main core of Ricoeur's hermeneutics[3] are a series of articles shattered on more than thirteen years. This simple fact has an undeniable influence on one's proper understanding of Ricoeur's hermeneutics. Even more, texts as such cannot be viewed as coherent sets of texts answering to a unique problem. Every text has different purposes, different circumstances, different aims, etc. Thus, when we approach Ricoeur's hermeneutics, a methodological decision is required. The critical moment in his hermeneutics (by critical moment we understand the moment in which Ricoeur becomes involved in what we can call a proper hermeneutic debate), and we have to emphasize this as clear as possible, is to be found in no text before the volume *Du texte à l'action* (1986).

The volume *Le conflit des interprétations* (1969), despite of its pronounced hermeneutic character, cannot be viewed as his hermeneutic theory, or, more proper, this volume is (1) neither developing a strict set of rules as any technique of interpretation requires, nor (2) debating on the status of understanding the understanding. With the exception of one article ("Existence et Herméneutique," on which I am going to draw later on, in order to prove its non-hermeneutic intents), Ricoeur is not at all interested neither in point 1, nor 2. Here we have what Ricoeur will later call 'symbolic hermeneutics,'[4] a hermeneutics axed on the double meaning of symbols, more precisely, on the second level of denotation of every symbolical thing. The well-known axiom "the symbol gives rise to thought"[5] will mark entirely his work until the volume *Du texte à l'action*. This is to say that Ricoeur is using hermeneutics as 'hermeneutic intelligence,' a sort of reflexive philosophy, in order to approach the hidden meaning of every symbol, or, in order to approach what symbols raised i.e., something to think about. It is not our purposes to follow here the development of Ricoeur's hermeneutics; it is enough if we state that Ricoeur's hermeneutics develops solitary without any connection with Gadamer's hermeneutics. The inner need to develop a hermeneutics responds to the problem of evil: "L'herméneutique du mal n'est pas une province

[2] Jean Grondin, *L'universalite de l' herméneutique* (Paris: PUF, 1993), XIV.

[3] Paul Ricoeur, *Le conflit des interprétations: Essais d'herméneutique* (Paris: Seuil, 1969), and *Du texte a l'action* (Paris: Seuil, 1986).

[4] See Ricoeur, *Du texte à l'action*, 30.

[5] See Paul Ricoeur, *Philosophie de la volonté. Tome II. Finitude et culpabilité* (Paris: Aubier, 1960), 479.

indifférente, mais la plus signifiante, peut-être *le lieu de naissance du problème herméneutique.*"[6] Nevertheless, with the volume *Du texte à l'action* things radically change.

Starting from this volume we want to interrogate the epistemological and, at the same time, the ontological status of Ricoeur hermeneutics.

In a series of articles Ricoeur asks himself if it wouldn't be more plausible to call the fundamental work of Gadamer *Truth OR Method* and not *Truth AND Method.*[7] Of course, here we have, among other critical approaches of Gadamer's philosophy to which we are going to come back later in our study, the fundamental charge that Ricoeur brings on Gadamer. That is to say, in more or less Ricoeur's words, that Gadamer's hermeneutics is too ontological and it almost breaks the ties with epistemology. Choosing between *alienated distanciation* and *appropriated experience* is the fundamental error (*aporia*) of the gadamerian hermeneutics. If, on the one side, we are going to follow the path opened by epistemology, we must give up our proximity to truth; if, on the other side, we remain on the level of *appropriated experience,* i.e., ontology, we are losing from sight the methodological aspect. This is the reason why Ricoeur urges us to understand that Gadamer should name his work *Truth OR Method*. The disjunction *OR* emphasises better the brake between ontology and epistemology. The conjunction *AND* presupposes a sort of equivalence between ontology and epistemology, an equivalence that is, in Ricoeur's eyes, nowhere to be found in Gadamer's work.

Thus, starting from this point we want to address the same question to Ricoeur. Although we cannot find in Ricoeur's work, like in Gadamer's, a single book (and as simple as it might appear this fact creates an on-going debate among exegetes) that is treating the hermeneutic problem, we have nevertheless a compact series of articles that are treating in a coherent manner the hermeneutic problem. I want to emphasise, as strongly as possible, that, here, at the level of *Du texte a l'action* we have the main core of Ricoeur's hermeneutics. Here we will find his hermeneutic debate with Schleiermacher, Dilthey, Heidegger and Gadamer. Here we have his hermeneutic theory from where we can talk, in the strong sense of the word, about Ricoeur's hermeneutics. Here we will find his criticism to the hermeneutic tradition, and, once again, his version of hermeneutics. But, in the same time, here as well, we will observe his non-hermeneutic intents.

---

[6] "The hermeneutics of evil is not an indifferent region but the most significant one, perhaps, *the birth place of the hermeneutic problem in itself* ." (Ricoeur, *Le conflit des interprétations,* 313, my translation & my emphasis.)

[7] See Ricoeur, *Du texte à l'action,* 97, 101. For the same purposes see his *Écrits et conférences 2: Herméneutique*, (Paris: Seuil, 2010), 134.

Hence, starting from the middle of Ricoeur's hermeneutics, we can deploy our question, our approach and our answer. The question: how can be regarded Ricoeur's hermeneutics: as an ontology or as an epistemology? Our approach: from the middle of the volume *Du texte à l'action* moving backwards to his work where hermeneutics is strongly tied to ontology, and forward to *Temps et récit,* where hermeneutics, in spite of its epistemological character, is still linked to ontology. Our answer: even if Ricoeur denied or, better, refused the ontological nature of hermeneutics, he is proposing a new hermeneutic ontology or, in a more suitable manner, a reversed ontology.

2. Hermeneutics between the *Cogito* and understanding

The critical approach of what we know today as *philosophical hermeneutics* in Ricoeur's case does not start with Gadamer, but with Heidegger. Even if we can agree that the epistemological moment of Ricoeur's hermeneutics can be found in a text like "Existence et Herméneutique," I want to establish that the main purpose of Ricoeur's argumentation is not related entirely to hermeneutics, but to anthropological philosophy. But before that, I will make a short comment on the role of structuralism in the Ricoeur's thought regarding hermeneutics. We will find that structuralism plays a double role in Ricoeur's thought: firstly as an opponent and later as an ally. I think this is important due to the fact that his critical hermeneutics, or more properly, his epistemological hermeneutics, relies entirely on structuralism. In a text like "Herméneutique et structuralisme," Ricoeur debates on the status of a structural approach to "a general theory of sense" (*théorie générale du sens*).[8] What we have to emphasize here is the fact that Ricoeur is using 'hermeneutic intelligence' (*intelligence herméneutique*) [9] as a valid method to enter the world of 'sense.' The failure of structuralism is that it cannot explain the 'surplus of sense'[10] that we find in any myth or symbol. This is the reason why we need a hermeneutic intelligence: so that we should be able to access the surplus of meaning. His argumentation follows in a critical manner an on-going debate on the temporal status of myth (diachronic or synchronic) emphasizing the need to adopt another view (other than structuralism) when we deal with the diachronic status of myth. It is important to stress here the refusal of structuralism. Of course, Ricoeur is not denying any relevance to structuralism[11] but, in 1963, for him, hermeneutic intelligence is the proper way to enter the 'world of sense.' In other words, it is

[8] Ricoeur, *Le conflit des interprétations*, 42.
[9] Ricoeur, *Le conflit des interprétations*, 48.
[10] Ricoeur, *Le conflit des interprétations*, 51.
[11] Ricoeur, *Le conflit des interprétations*, 57.

about a clash: *hermeneutic intelligence* against *structuralism*. We will draw later on the positive role that structuralism plays in his critical hermeneutics.

Hence, in "Existence et Herméneutique," we find for the first time Ricoeur's disapproval on Heidegger's direct ontology of understanding. We cannot follow the path opened before us by Heidegger, Ricoeur tells us, because we are going to lose from our sight the methodological features of hermeneutics.[12] We have to maintain a close contact with the methodological trains of hermeneutics if one wishes to establish a valid set of rules of the human sciences, especially the historical sciences.[13] In other words, and much closer to Ricoeur's debate, one cannot follow Heidegger's ontology of understanding if he wishes to fully understand the conflict of interpretation.[14] It seems that in 1965, when the article was written, Ricoeur is emerging himself without any doubt in the classical problems of hermeneutics regarding the problem of understanding as such, and the problem of a methodology required by the human sciences. But if we are going to closely follow his entire argumentation, we can notice a slight change of subject. The entire second part of the article is concerned on what is known as *anthropological philosophy*. His main interest is the *Cogito* as a debatable problem for structuralism, psychoanalysis and phenomenology of religion. Appealing to the 'semantic knot,'[15] or in a more proper way, to the symbol, as the center of all hermeneutics particular or general,[16] Ricoeur denies any attempt to knowledge that starts from the center of the human subject. We need, in his words, an archaeology of the subject and not an ontology of understanding.[17] We can never begin from the center of the *Cogito,* but always from the long route of symbols. *Thus, Ricoeur's main objection to Heidegger is not, even if it seems so, the problem of understanding as such, but the problem of the Cogito, the problem of understanding the Cogito*. It is true that we have here an intercalated problem, but for Ricoeur the problem to solve is an anthropological one and not a hermeneutic one. But, for the moment, it is important to notice the double task of Ricoeur's analysis: hermeneutic at the beginning, anthropological at the end. Better yet, so as to fully understand the anthropological task of his endeavor, in Ricoeur's eyes the ontology of understanding proper to Heidegger is similar to the vain pretention of Descartes' *cogito*.[18] Because reflexion (as

[12] Ricoeur, *Le conflit des interprétations*, 11.
[13] Ricoeur, *Le conflit des interprétations*, 14.
[14] Ricoeur, *Le conflit des interprétations*, 14.
[15] Ricoeur, *Le conflit des interprétations*, 16.
[16] Ricoeur, *Le conflit des interprétations*, 14.
[17] Ricoeur, *Le conflit des interprétations*, 25.
[18] Ricoeur, *Le conflit des interprétations*, 21.

understanding or as *cogito*) always is 'our effort to exist;'[19] this is to say, reflexion is never given but always acquired. What we have here is a double critique: firstly towards the pretention of knowledge of the Cartesian *cogito*, and, secondly, to the pretention of a direct ontology of understanding. We always must have a detour, a route, a path to follow, so that we should to be able to understand the *cogito* or the understanding. Ricoeur's interest is not the problem of understanding, but the problem of the *self* which knows himself and discovers himself. Bluntly, Ricoeur's main interest is the construction of the self and not the construction of understanding.

Of course, we cannot deny and we are not denying the hermeneutic debate offered by Ricoeur in this article. The fact is that, indeed, Ricoeur is defending here the epistemological status of hermeneutics, asking to maintain a close contact to the main trains of hermeneutics, understood as methodology, but the core of his argumentation is not to offer us the 'dreamed' methodology of the human sciences, but to argue against a direct path to the *cogito*. Again, we have to insist on the fact that for Ricoeur, at the level of *Le conflit des interpretations,* hermeneutics, understood as hermeneutic intelligence, plays an almost indefinable role. With hermeneutic intelligence we are able to enter the world of the sense, beyond structuralism or psychoanalysis.[20] But, in the same time, when we adopt an ontological position, such as Heidegger's, we have to refuse the direct access to understanding (because it is similar to the direct access to the *cogito*) only to propose an alternative way, a long route, an indirect ontology. This indirect path to the problem of *cogito* will have a dramatic influence on what we can call his hermeneutic theory.

3. Hermeneutics as/or reversed ontology

We have to assert that Gadamer plays a fundamental role on the development of Ricoeur's critical hermeneutics. Despite the fact that Ricoeur acknowledges that his hermeneutics develops 'step by step,'[21] always discovering another problematic aspect of hermeneutics, and despite the fact that his initial encounter with hermeneutics has no connections with Gadamer's, the critical moment in Ricoeur's hermeneutics is, undoubtedly, an on-going debate with Gadamer. We are not going to deal here with all the problems raised by this fact; even more, we are not

[19] "La réflexion est l'appropriation de notre effort pour exister et de notre désir d'être a travers les œuvres qui témoignent de cet effort et de ce désir." (The reflexion is our effort to exist and our desire to be through the works that confine this effort and this desire.") (Ricoeur, *Le conflit des interprétations*, 21, my translation.)

[20] See Ricoeur, *Le conflit des interprétations*, 101-195.

[21] Ricoeur, *Écrits et conférences 2*, 17.

interested in establishing the inner connections between Gadamer and Ricoeur or the main differences between them. However, in order to grasp the indirect ontology that Ricoeur is proposing, we cannot exclude Gadamer. Obviously, Ricoeur's ontology does not fundamentally change with his encounter with Gadamer, but it receives a new insight, a new determination. Thus, his commitment to the 'long route' into the world of symbols, monuments and written works in order to have access to the *cogito*[22] or, into the world of understanding, is kept. Even if the debate is moved from the *cogito* to *understanding*, or even if Ricoeur is fully committed to a hermeneutic debate, his non-hermeneutic insights regarding the possibility of knowing the self remains the same.

The fact is that his engagement to the epistemological characteristics of hermeneutics explodes in a critical manner in three consecutive texts, that we can call the center of Ricoeur's critical hermeneutics.[23] Thus, in "La tache de l'herméneutique," Ricoeur is renewing his commitment to the epistemological insights of hermeneutics, accusing Gadamer that he is too close to the claims of Heidegger. In other words, as we saw, that his hermeneutics is too ontological. In order to be able to have an epistemological hermeneutics, the first thing that we must do is to restrict hermeneutics to the text. After all, the text is the main center of all hermeneutics.[24] Second, we have to give the text a sort of full autonomy, so as that we should to be able to approach it as strictly as possible: autonomy regarding its author, autonomy regarding its destination, autonomy regarding its socio-economic factors.[25] This triple autonomy confines the text with the much needed objectivity. Freed from its philological, historical and sociological elements, the text presents itself in front of us as an autonomous object. Our hermeneutic effort or, more precisely, our epistemological effort will respect this autonomy of the text by applying the fundamental set of laws offered by structuralism. We do not have to borrow the methods from the natural sciences, but we have to look within our own (humanistic) field towards structuralism. Thus, hermeneutics will be able to adopt an epistemological status.

---

[22] See Ricoeur, *Le conflit des interprétations,* 10. The same statement appears also in *Du texte a l'action,* 116.

[23] The first critical text in which Ricoeur is proposing a 'new concept of interpretation' – "Qu'est-ce qu'un texte?" – is written in 1970. Together with "La tache de l'herméneutique," "La fonction hermeneutique de la distanciation," and "Hermeneutique philosophique et hermeneutique biblique," all written in 1975, "Qu'est-ce qu'un texte?" can be viewed as his theoretical and critical hermeneutics.

[24] Ricoeur, *Du texte a l'action,* 75.

[25] Ricoeur, *Du texte a l'action,* 111,112, 114, 116.

Here structuralism plays a positive role. If, for instance, as we saw, confronting structuralism Ricoeur is proposing hermeneutic intelligence as a valid method in order to enter the world of the sense, confronting ontological hermeneutics, structuralism functions as a valid method in order to offer the text a much needed autonomy. Bluntly, we can find a dialectical movement within Ricoeur's hermeneutics: the refusal of structuralism (the impossibility to accede to the sense) as a first movement, but, in a second movement, the refusal of ontological hermeneutics as a vain pretention to have access to the sense. Of course, we have here two very different utilizations of the same concept: firstly an anthropological usage of hermeneutics and secondly a proper hermeneutic usage of hermeneutics.

Evidently the refusal of a direct ontology assumes as a counterweight a clear methodological position. But a peculiar position that, at the end, will become an ontological position. What we are arguing is that Ricoeur, in his debate with Gadamer, has a similar position, as that assumed in his debate with Heidegger. In other words his anthropological position is to be found in his hermeneutic position. That is to say the refusal of Gadamer's ontology implies the proposal of another type of ontology.

Thus, we have to follow closely Ricoeur's argumentation against Gadamer. The fact is that Ricoeur is engaging in a direct debate with Gadamer's main concepts: *distinction alienante* (*Verfremdung*) and *raport primordial d'appartenance* (*Zugehorigkeit*).[26] Therefore, if for Gadamer *Verfremdung* makes one to lose the contact with the things to be understood for Ricoeur, the same *Verfremdung* has a positive role.[27] To give an example, perhaps the most important one, historical conscience enrolls in the same time *Verfremdung* and *Zugehorigkeit*; the need to understand a historical fact is possible due to the fact that I am a historical being (preceded by historical facts), but the impossibility to understand the fact is given by the temporal distance. How can we understand critically a historical fact if we are preceded by the fact itself? Or, better, how can we understand ourselves between *Verfremdung* and *Zugehorigkeit*? In Ricoeur's eyes, *Zugehorigkeit* with its ontological structure obstructs the critical position of hermeneutics.[28] *Zugehorigkeit* can never be abolished but also its negative character cannot be abolished. But in this manner we will never be able to introduce a critical hermeneutic position. This is why Ricoeur argues for a more positive understanding of *Verfremdung*.

[26] Ricoeur, *Du texte à l'action,* 96.

[27] Ricoeur, *Du texte à l'action,* 117.

[28] Ricoeur, *Du texte a l'action,* 101- 102. See also his "Herméneutique et critique des idéologies," in *Du texte a l'action,* 333-378.

Yet, we may ask, how this positive understanding of *Verfremdung* can be achieved? Ricoeur's answer is: by appealing to the text, or, more clearly, to the *world of the text*. The work of the text is able to defeat *Verfremdung* in which our understanding loses itself.[29] What we must understand is that Ricoeur offers to *Verfremdung* a positive character. In this way, historical distance does not become an impediment to our understanding. But what exactly means positive side of *Verfremdung*? In fact, what we have here is a shift of view. In the center of our act of comprehension we are not going to find the human subject with his constitutive prejudices but the self, formed and informed by the text. The world opened by the text, its work, shapes the human subject. In fact, the world of the text is the center of all comprehension. The work of the text which shapes my conscience is the center of all hermeneutics. We can here observe better why Ricoeur stipulated the full autonomy of the text. In interpreting a text we are not searching for the author, we are not searching for the hidden intention, we are not searching to understand the text better than the author himself, we are not searching for its socio-economic factors, but we are searching for the *sense* of the text which forms the self. In this moment, a critical question may be introduced. How can we control in an epistemological manner the world of the text? The internal structure of Ricoeur's epistemology may at the end be reduced to a simple and single operation; to be affected by the text, to find yourself in front of the text. However, isn't it safe to affirm that in this manner we are losing the contact with epistemology and we enter a different type of ontology?

We can restrict the entire tension created between epistemological and ontological implications of Ricoeur's hermeneutics to the following citation:

[29] Moreover, in this way hermeneutics keeps a close contact with its epistemological trends. We have to understand that, in Ricoeur's eyes, the experience of *Zugehorigkeit* still holds to a philosophy of cogito. See for instance his critique towards Gadamer that he did not overcome the romantic status of hermeneutics (*Du texte a l'action,* 97). Of course, we can observe here the same critique as in the case of Heidegger: abolishing the pretention of a foundational cogito. See *Du texte à l'action,* 52, where Ricoeur asserts: "*Une manière radicale de mettre en question le primat de la subjectivité est de prendre pour axe hermenéutique la théorie du texte. Dans la mesure où le sense d'un texte s'est rendu autonome par rapport à l'intention subjective de son auteur, la question esentielle n'est pas de retrouver, derrière le texte, l'intention perdue, mais de déployer, devant le texte, le 'monde' qu'il ouvre et découvre.*" ("*A radical way of placing the primacy of subjectivity in question is to take the theory of the text as the hermeneutic axis. Insofar as the meaning of a text is rendered autonomous with respect to the subjective intention of its author, the essential question is, not to recover, behind the text, the lost intention but to unfold, in front of the text, the world it opens up and discloses.*") (my translation, emphasis original.)

> Non point imposer au texte sa propre capacité finie de comprendre, mais s'exposer au texte et recevoir de lui un soi plus vaste, qui serait la proposition d'existence répondant de la manière la plus appropriée à la proposition de monde … *La compréhension est alors tout contraire d'une constitution dont le sujet aurait la clé. Il serait à cet égard plus juste de dire que le soi est constitué par la 'chose' du texte*.[30]

Not only the understanding of the text as an object of hermeneutics is advanced here, but, more important, the text itself has direct implications in understanding as such. This position can be confusing for one who is searching for the epistemological fundaments of Ricoeur's hermeneutics. 'The world of the text' or 'the work of the text,' overturns the connections between epistemology and ontology. Thus, the proposition of sense specific to the 'world of the text' cannot be eradicated into an epistemological endeavor. The ontological status is here evident: the sense is pre-given, there is a world of the text that informs my conscience, but exactly this proposition of sense cannot be epistemologically analysed. The sense of the text, its world, becomes the center of the hermeneutic experience specific to Ricoeur's theory.

We can at this moment draw a general conclusion. *Verfremdung* can be viewed in its positive aspect due to the fact that it produces a shift from our pre-judgments to the world of the text. Or, more clearly, from the *cogito* to the text. But, in this way, the autonomy of the text has a relative, if we can put it like this, autonomy; until the moment we have to accept that the sense of the text has a direct implication to my understanding. At this point, any epistemological endeavor cannot be produced. In front of a direct ontology of understanding, as we can find at Heidegger and Gadamer, a reverse ontology is asserted. We are not going to start from the inner conscience of the self but from the *sense* outside of it. We are not going to start from the pre-comprehension but from the center of the *sense* of a text. Hermeneutics is the theory which has its main tasks to discover the proposition of sense given by any text into the conscience of the self.[31] A last

[30] „We cannot impose to the text one's own capacity of understanding but exposing himself to the text and receiving back a self more vast ... *Understanding is then divergent to a constitution where the subject is the key. It is, then, more correctly to say that the self is constituted by the 'work' of the text.*"(Ricoeur, *Du texte à l'action,* 117, my translation & emphasis).

[31] 'Narrative identity' as the central concept of *Temps et récit* will elaborate on this, as we called it, reversed ontology. It was beyond my intention to follow Ricoeur's narrative identity. Even more, it was not my intention to follow here the inner connection between the theory of text, action and history in Ricoeur's thought. For a critical approach see Jean Grondin, "L'herméneutique positive de Paul Ricoeur. Du temps au récit," in *Temps et récit de Paul Ricœur en débat*, ed. Christian Bouchindhomme and Rainer Rochlitz (Paris: Cerf, 1990), 111-121. For the

remark is required. The problem of understanding is for Ricoeur, even if the fact as such is not declared, similar to the problem of *cogito*. How can one understand the *cogito* or how one can understand understanding – these are, in Ricoeur's eyes, similar questions. If in the first place we will never begin from the vain pretentions of the *cogito*, in the second instance, we will never begin from the core of understanding. Neither understanding nor cogito stand as our departing points. On the contrary, we always arrive at them, never begin with them. Nevertheless, if one strives to find Ricoeur's ontology he will have to understand that:

> "L'ontologie est bien la terre promise pour une philosophie qui commence par le langage et par la réflexion; mais, comme Moise, le sujet parlant et réfléchissant peut seulement l'apercevoir avant de mourir."[32]

epistemological or ontological status of narrative identity, see David Carr, "Epistemologie et ontologie du récit," in *Paul Ricoeur. Les métamorphoses de la raison herméneutique*, ed. Jean Greisch and Richard Kearney (Paris: Cerf, 1991), 205-214.

[32] „The ontology is the Promised Land for a philosophy that begins with language and with reflexion; but, like Moses, the human subject can only perceive it before dying." (Ricoeur, *Le conflit des interprétations*, 28, my translation.)

A SIMPLE SOLUTION TO THE TWO ENVELOPE PROBLEM

Ned MARKOSIAN

ABSTRACT: Various proposals have been made for solving The Two Envelope Problem. But even though the problem itself is easily stated and quite simple, the proposed solutions have not been. Some involve calculus, some involve considerations about infinite values, and some are complicated in other ways. Moreover, there is not yet any one solution that is widely accepted as correct. In addition to being notable for its simplicity and its lack of a generally agreed-upon solution, The Two Envelope Problem is also notable because it demonstrates that something that has been taken to be a fundamental principle of decision theory is false. The main purpose of this paper is to propose and defend a simple solution to The Two Envelope Problem. But I also want to make a start on the project of figuring out how the relevant fundamental principle of decision theory should be revised.

KEYWORDS: two envelope problem, decision theory, expected utility, decision under risk

1. The Problem

The Two Envelope Problem concerns a game – The Two Envelope Game – that goes like this. First, two seemingly indistinguishable envelopes, A and B, are presented to you. Each envelope contains some finite, non-zero amount of money, in dollars, and one envelope contains twice as much money as the other. Moreover, these facts about A and B are known to you. But you do not know how much money envelope A contains, or how much money envelope B contains, or even the total amount in the two envelopes. In this initial stage of the game, you are required to choose one of the envelopes, which then becomes your personal property. Then, in the second stage of the game, you are given the opportunity to trade your envelope for the other envelope. After the second stage of the game, you get to keep whichever envelope you are left with (together with its contents).

That's the whole game. It is a simple one, and a good one to play. Either way you are going to make some money.

Now, commonsense tells us that when you get to the second stage of The Two Envelope Game – where you get to decide whether to trade the envelope you have initially chosen for the other envelope – there is no special reason why you

© LOGOS & EPISTEME, II, 3 (2011): 347-357

should trade. After all, given your evidence, you are just as likely to have chosen the envelope with more money as you are to have chosen the envelope with less money. And so, although it seems clear that you are rationally permitted to trade, in the second stage of The Two Envelope Game, it also seems clear that you are rationally permitted to stick with the envelope you initially chose.

But it turns out that there is an apparently sound argument, based on a fundamental and seemingly uncontroversial principle of decision theory, for the conclusion that it is rationally obligatory to trade in The Two Envelope Game.

The relevant principle of decision theory concerns the correct way to calculate expected utilities in what is known as a 'decision under risk' situation. A decision under risk situation is one in which the agent does not know with certainty what the outcomes of the relevant alternatives would be, but does know, for each outcome that might result from a given action, the *probability* of that outcome's resulting from that action, as well as the *value* of that outcome. Here is the principle.

> (FP) For any alternative, A, available to an agent, S, the *expected utility* of A can be calculated in the following way. First, determine all of the possible outcomes that might result from S's doing A. Second, for each such outcome, determine both the *probability* of that outcome (given that S does A) and the *value* of that outcome. Third, for each outcome, multiply the relevant probability times the relevant value. And finally, calculate the sum of all of these products. That sum = the expected utility of A.[1]

FP is meant to be one of the fundamental principles of decision theory (thus the abbreviation "FP"). It might at first appear to be complicated and somewhat arcane, but the idea that FP captures is actually very simple and intuitive. It would be hard to deny that this idea is on some level behind much of the reasoning that we (correctly) employ in everyday decision-making.

In any case, here is the argument from FP to the conclusion that it is rationally obligatory to trade envelopes when you get to the second stage of The Two Envelope Game. Let n = the amount of money in your envelope. Then, since there is a .5 probability that trading will mean trading up (thereby doubling your money), and a .5 probability that trading will mean trading down (thereby halving

[1] For principles relevantly like FP, see, for example, Brian Skyrms, *Choice and Chance: An Introduction to Inductive Logic,* 4th ed. (Belmont: Wadsworth, 2000), 128; and Michael D. Resnick, *Choices: An Introduction to Decision Theory* (Minneapolis: University of Minnesota Press, 1987), Ch. 3.

your money), FP entails that we can calculate the expected utility of trading as follows.[2]

expected utility of trading = $[(.5 \times 2n) + (.5 \times .5n)] = 1.25n$

Meanwhile, it is clear that if you stick, then you are guaranteed to end up with n (since that is the amount in your envelope). So FP entails that we can calculate the expected utility of sticking as follows.

expected utility of sticking = $(1 \times n) = n$

Given these expected utilities (and assuming that it is always rationally obligatory to maximize expected utility), it looks like it is rationally obligatory to trade.

Since we know from commonsense that it is not rationally obligatory to trade in The Two Envelope Game, our problem is to say what, exactly, is wrong with this plausible argument.

Various proposals have been made for solving The Two Envelope Problem.[3] But, strangely enough, even though the problem itself is easily stated and quite simple, the proposed solutions have not been. Some involve calculus, some involve considerations about infinite values, and some are complicated in other ways. Moreover, there is not yet any one solution that is widely accepted as correct. My conviction is that a problem as simple as The Two Envelope Problem must have a simple solution. (And by 'a simple solution' I mean, roughly, one that does not involve any math more complicated than algebra.)

---

[2] Assuming, for the sake of simplicity, that the expected utility of any alternative = the expected dollar utility of that alternative.

[3] For more on the problem, see, for example, Barry Nalebuff, "The Other Person's Envelope Is Always Greener;" *Journal of Economic Perspectives* 3 (1989): 171-181; Frank Jackson, Peter Menzies, and Graham Oppy, "The Two Envelope 'Paradox,'" *Analysis* 54 (1994): 43-45; Paul Castel and Diderik Batens, "The Two-envelope Paradox: The Infinite Case," *Analysis* 54 (1994): 46-49; John Broome, "The Two-envelope Paradox," *Analysis* 55 (1995): 6-11; Timothy McGrew, David Shier, and Harry Silverstein, "The Two-envelope Paradox Resolved," *Analysis* 57 (1997): 28-33; Alexander D. Scott and Michael Scott, "What's in the Two Envelope Paradox?" *Analysis* 57 (1997): 34-41; Frank Arntzenius and David McCarthy, "The Two Envelope Paradox and Infinite Expectations," *Analysis* 57 (1997): 42-50; Michael Clark and Nicholas Shackel, "The Two-envelope Paradox," *Mind* 109 (2000): 415-442; Terry Horgan, "The Two-Envelope Paradox, Nonstandard Expected Utility, and the Intensionality of Probability," *Nous* 34 (2000): 578-603; David Chalmers, "The St. Petersburg Two-envelope Paradox," *Analysis* 62 (2002): 155-157; James Chase, "The Non-probabilistic Two Envelope Paradox," *Analysis* 62 (2002): 157-160; and Franz Dietrich and Christian List, "The Two-Envelope Pradox: An Axiomatic Approach," *Mind* 114 (2005): 239-248.

In addition to being notable for its simplicity and its lack of a generally agreed-upon solution, The Two Envelope Problem is also notable because it demonstrates that something that has been taken to be a fundamental principle of decision theory – namely, FP – is false. The main purpose of this paper is to propose and defend a simple solution to The Two Envelope Problem. But I also want to make a start on the project of figuring out how FP should be revised.

2. A Variety of Ways to Calculate Expected Utilities in The Two Envelope Game

Notice that we could just as easily reason about the expected utility of trading in The Two Envelope Game as follows. Let m = the amount of money in the other envelope. Then we get these calculations.

expected utility of trading = $(1 \times m) = m$

expected utility of sticking = $[(.5 \times 2m) + (.5 \times .5m)] = 1.25m$

Given this way of calculating expected utilities, it looks like it is rationally obligatory to stick.

Or we could reason as follows. Let x = the lesser of the two amounts in the envelopes. Then if you trade, there are two possibilities: either you will have x in your envelope and will double your money by trading, or you will have 2x in your envelope and will halve your money by trading. Whereas if you stick, then either you will be sticking with x or you will be sticking with 2x. Thus, we get these calculations.

expected utility of trading = $[(.5 \times 2x) + (.5 \times x)] = 1.5x$

expected utility of sticking = $[(.5 \times x) + (.5 \times 2x)] = 1.5x$

According to this way of calculating expected utilities it is neither rationally obligatory to trade nor rationally obligatory to stick.

There's more. We could also reason as follows. Let y = the greater of the two amounts in the envelopes. Then if you trade, there are two possibilities: either you will have y in your envelope and will halve your money by trading, or you will have .5y in your envelope and will double your money by trading. Whereas if you stick, then either you will be sticking with y or you will be sticking with .5y. We thus get these calculations.

expected utility of trading = $[(.5 \times .5y) + (.5 \times y)] = .75y$

expected utility of sticking = $[(.5 \times y) + (.5 \times .5y)] = .75y$

And according to these calculations, it is neither rationally obligatory to trade nor rationally obligatory to stick.

And finally, since there is one other significant quantity in the game that we could name, we could reason like this. Let z = the total amount in the two envelopes. Then one envelope contains 1/3z, while the other envelope contains 2/3z, so that if you trade, there is a .5 probability that you will trade up from 1/3z to 2/3z, and a .5 probability that you will trade down from 2/3z to 1/3z. Whereas if you stick, there is a .5 probability that you will be sticking with 1/3z and a .5 probability that you will be sticking with 2/3z. Thus, we get these calculations.

expected utility of trading = [(.5 × 1/3z) + (.5 × 2/3z)] = .5z

expected utility of sticking = [(.5 × 1/3z) + (.5 × 2/3z)] = .5z

According to this last way of calculating expected utilities, it is again neither rationally obligatory to trade nor rationally obligatory to stick.

We have seen that there are 5 important amounts of money involved in The Two Envelope Game that we can identify, and that, for each one, there is a corresponding way of calculating the expected utilities of trading and sticking.[4] These five ways can be summarized as follows.

Wn n = the amount in your envelope. It is rationally obligatory to trade.

Wm m = the amount in the other envelope. It is rationally obligatory to stick.

Wx x = the lesser of the two amounts in the envelopes. It is neither rationally obligatory to trade nor rationally obligatory to stick.

Wy y = the greater of the two amounts in the envelopes. It is neither rationally obligatory to trade nor rationally obligatory to stick.

Wz z = the total amount in the two envelopes. It is neither rationally obligatory to trade nor rationally obligatory to stick.

[4] More accurately, there are five ways we can pick out an important amount of money involved in The Two Envelope Game – some pairs of which pick out the same amount of money (since, for example, the amount in your envelope may be the same amount as the lesser of the two amounts in the envelopes) – such that for each one, there is a corresponding way of calculating the expected utilities of trading and sticking.

So what should we do? Each of the five ways of calculating the relevant expected utilities seems to have a legitimate claim to being correct. Yet they yield results that are jointly inconsistent. Should we just observe that three of the ways yield the same result (that it is neither rationally obligatory to trade nor rationally obligatory to stick), while the other two 'cancel each other out,' and conclude that it is neither rationally obligatory to trade nor rationally obligatory to stick? Although that approach would give us the answer that we intuitively know to be correct, it doesn't really seem to be the right method for dealing with this kind of problem. Moreover, it will be recalled that our problem is not to show that, in The Two Envelope Game, it is neither rationally obligatory to trade nor rationally obligatory to stick, but, rather, to say what is wrong with the line of reasoning for the conclusion that it is rationally obligatory to trade in The Two Envelope Game.[5] In other words, what we need is some independent reason for claiming that Wn and Wm are wrong, while Wx-Wz are right. But is there such a reason?

3. Missing Information

I think there is. In particular, I think it can be shown that Wn and Wm are both incorrect ways of calculating the relevant expected utilities, and that Wx, Wy and Wz are all correct. The reason for this is that there is a certain crucial piece of information available to us about The Two Envelope Game that Wn and Wm (but not Wx-Wz) fail to take into account. Allow me to explain.

Consider a game that I will call *Double or Half*, or *DOH*. DOH is like the more traditional Double or Nothing, in that you start with some particular (finite, non-zero) sum of money, and you may choose either to play or not to play. If you decide to play, then some apparently indeterministic procedure with two possible outcomes that seem equally likely – such as a coin flip with a fair coin – is used to determine whether you win or lose. If you win, then your money is doubled, and if you lose, then your money is halved.

As any competent gambler will tell you, you should always jump at the chance to play DOH. For in DOH, the odds of winning are equal to the odds of losing, but you stand to win more than you stand to lose. I.e., if we let n = the amount of money you start with, then the following (familiar) calculations give us the expected utility of playing DOH and the expected utility of not playing DOH.

expected utility of playing DOH = $[(.5 \times 2n) + (.5 \times .5n)] = 1.25n$

expected utility of not playing DOH = $(1 \times n) = n$

[5] Presumably whatever flaw we identify in Wn will also appear in Wm.

In light of these expected utility calculations, we can see that the following principle is a good one.[6]

PlayDOH: It is always rationally obligatory to play DOH.

Now consider another game, which I will call *Thirds*. In Thirds, some particular amount of money is first selected, without your knowing what it is, and then that amount is divided into two smaller amounts, so that one smaller amount equals one-third of the total and the other smaller amount equals two-thirds of the total. Next comes the seemingly indeterministic part, which we will again suppose consists of the flipping of a fair coin. The coin flip in Thirds determines which of the two smaller amounts – one-third or two-thirds – is initially assigned to you. After that has been determined (but before you know whether you have been assigned one-third or two-thirds of the total), you get to decide whether or not you will trade your amount for the other amount. If you choose to trade, then you are choosing to play Thirds, and if you choose not to trade, then you are opting not to play Thirds.

If we let z = the total amount of money in the game, then here are the (again already familiar) expected utility calculations for playing Thirds.

expected utility of playing Thirds = $[(.5 \times 1/3z) + (.5 \times 2/3z)] = .5z$

expected utility of not playing Thirds = $[(.5 \times 1/3z) + (.5 \times 2/3z)] = .5z$

As these calculations show, it is neither rationally obligatory to play Thirds nor rationally obligatory not to play Thirds.

Now let's return to our Two Envelope Problem and the five ways of calculating expected utilities in The Two Envelope Game. Consider Wn, the original way of calculating expected utilities in The Two Envelope Game (and the way that was involved in the argument for the conclusion that it is rationally obligatory to trade). It should be clear that Wn is based on the assumption that when you trade in The Two Envelope Game, you are playing DOH. But this assumption is false. To trade in The Two Envelope Game is not to play DOH, because when you trade in The Two Envelope Game, it's not the case that, for some particular amount of money, a seemingly indeterministic procedure determines whether you will get double that amount or half that amount. Instead, what happens is this. First, an apparently indeterministic procedure determines whether you will initially receive one-third or two-thirds of some specific amount of money. After that, you are permitted to trade your amount for whichever one of the

[6] Again assuming that it is always rationally obligatory to maximize expected utility, and that expected utility = expected dollar utility.

smaller amounts you do not have. In other words, when you trade in The Two Envelope Game, you are playing Thirds rather than playing DOH. And that is why it is neither rationally obligatory to trade nor rationally obligatory to stick in The Two Envelope Game.

These observations about The Two Envelope Game, DOH, and Thirds allow us to see on an intuitive level not only what is right about Wx, Wy, and Wz (namely, that they are based on the observation that to trade in The Two Envelope Game is to play Thirds), but also what is wrong with Wn (namely, that it is based on the mistaken assumption that to trade in The Two Envelope Game is to play a game of DOH).[7]

Unfortunately, however, the above observations about The Two Envelope Game, DOH, and Thirds do not allow us to state in any precise way exactly why Wn and Wm are incorrect. For there is nothing to stop the proponent of Wn, for example, from giving the following speech.

> I agree with everything you've said about The Two Envelope Game, DOH, and Thirds. In particular, I admit that to trade in The Two Envelope Game is *not* to play DOH. But I never said it was, nor did anything in my calculations. All my calculations are committed to is this undisputed truth: As far as you know, there is a .5 probability that if you trade in The Two Envelope Game then you will be trading up (thereby doubling your money), and a .5 probability that if you trade in The Two Envelope Game then you will be trading down (thereby halving your money). Since that is all my calculations entail, and since, in particular, my calculations do not entail anything false (such as that you would be playing DOH if you were to trade in The Two Envelope Game), there is nothing wrong with my calculations.

A proponent of Wn who said this would be right in her claim that, as far as you know, there is a .5 probability that if you trade in The Two Envelope Game then you will be trading up (thereby doubling your money), and a .5 probability that if you trade in The Two Envelope Game then you will be trading down (thereby halving your money). But she would be wrong in claiming that her calculations do not entail anything false. For our thinking about DOH and Thirds has prepared the way for us to show, in a more precise fashion, exactly what is wrong with Wn and Wm.

First, let us note an obvious and uncontroversial fact about expected utility calculations. They are calculations of *expected* utilities. That is, they are calculations concerning how valuable the outcomes of various choices can be

[7] Similarly, the above observations about The Two Envelope Game, DOH, and Thirds allow us to see on an intuitive level what is wrong with Wm (namely, that it is based on the mistaken assumption that to stick in The Two Envelope Game is to play a game of DOH).

expected to be, *given the evidence available to us*. For this reason, it is possible for a way of calculating expected utilities to be incorrect simply because it fails to take into account some relevant piece of information that is available to us.

This fact about expected utility calculations will play an important role in my account of what is wrong with Wn and Wm. But in order to spell out the rest of that account, I will need to introduce another important fact about The Two Envelope Game. And in order to do that we will first need to introduce three technical terms. Let's say that you are playing a *trading game* iff you are playing a game in which, at some point, you have an opportunity to trade a certain amount of money that is already in your possession for some other amount of money. And let's say that you are playing a *fixed-total* trading game iff you are playing a trading game in which the total amount of money in the game (i.e., the amount initially in your possession + the amount that you can trade for) is fixed before it is determined how much money is initially in your possession. Finally, let's say that you are playing a *variable-total* trading game iff you are playing a trading game in which the total amount of money in the game is not fixed at the time when it is determined how much money is initially in your possession. Thus, for instance, Thirds is a fixed-total trading game (since the total amount of money in the game is already fixed by the time you are initially assigned your envelope), but DOH is a variable-total trading game (since the total amount of money in the game is not determined until after you are assigned your initial amount (and even then it is not determined unless you decide to trade)).

Now, here is a general fact about fixed-total trading games that can be easily proved. In any such game, the expected utility of trading = the expected utility of sticking = one half of the total amount in the game. And here is the proof. Suppose you are playing some fixed-total trading game. For any such game there are, before you decide whether to trade, three numbers, i, j, and k, such that the amount of money initially in your possession = i dollars, the amount that you are allowed to trade for = j dollars, and the total amount of money in the game (i.e., i dollars + j dollars) = k dollars. (This follows from the definition of a fixed-total trading game.) Then, since you have no reason to believe that you have either one of the two smaller amounts (namely, i and j) rather than the other, the expected utility of sticking is $(.5 \times i) + (.5 \times j)$. But that equals $.5 \times (i + j)$, and $(i + j) = k$. So the expected utility of sticking = .5k. And the same is true with regard to the expected utility of trading.

This important result about fixed-total trading games can be stated as follows.

> (IR) For any fixed-total trading game, G, the expected utility of trading in G = the expected utility of sticking in G = one half of the total amount in G.

It is important to note that IR is a purely general result. It applies to any fixed-total trading game, regardless of the specific amounts involved in the game, and regardless of the ratio between those amounts. Thus, no other information that is added to the fact that you are playing a fixed-total trading game can possibly affect this result. (With the exception, of course, of information that requires you to revise your estimation that you are playing a fixed-total game. But no such information is available to you in the case of The Two Envelope Game.) Indeed, even if you were allowed to open your envelope before deciding whether to trade in The Two Envelope Game, whatever you learned by doing so would not change the fact that you are playing a fixed-total game. It is for this very reason that opening your envelope in The Two Envelope Game does not change the fact that is is neither rationally obligatory to trade nor rationally obligatory to stick. And the simple reason for this is that learning the amount in your envelope does not change the fact that IR applies to your situation.

Let's return to the five ways of calculating expected utilities for trading and sticking in The Two Envelope Game. We can now see, given both the generality and the truth of IR, that the fact that the total amount of money in The Two Envelope Game is fixed before you get to decide whether to trade is of the utmost importance. And now we can also see what is wrong with Wn and Wm. Although Wx, Wy, and Wz all entail that the total amount in The Two Envelope Game is fixed (for Wx entails that the total amount = 3x, Wy entails that the total amount = 1.5y, and Wz entails that the total amount = z), neither Wn nor Wm does so (since all Wn entails about the total amount is that it is either 3n or 1.5n, and all Wm entails about the total amount is that it is either 3m or 1.5m).

The upshot is that Wn and Wm are defective, and give incorrect results, because they fail to incorporate a crucial fact about the situation in The Two Envelope Game that is available to us, namely, the simple fact that the total amount in The Two Envelope Game is fixed rather than variable.

4. Revising FP

Let's take stock. We have seen that FP – which has been taken to be one of the fundamental principles of decision theory – is in need of revision. For it can be used to yield false results in our situation involving The Two Envelope Game and, moreover, it yields inconsistent results in that situation. Meanwhile, I have argued that the problem with applying FP to our situation in The Two Envelope Game is that it is possible to calculate expected utilities in the way prescribed by FP while at the same time failing to take into account some information that is crucial to that situation.

These considerations suggest that we should simply add a clause to FP that explicitly says to take into account all relevant information, as in the following revised version of FP.

> (FP2) For any alternative, A, available to an agent, S, the *expected utility* of A can be calculated in the following way. First, determine all of the possible outcomes that might result from S's doing A. Second, for each such outcome, determine both the *probability* of that outcome (given that S does A) and the *value* of that outcome. *(But be sure to individuate these possible outcomes, probabilities, and values in such a way that no crucial information about S's situation is left out of the reckoning.)* Third, for each outcome, multiply the relevant probability times the relevant value. And finally, calculate the sum of all of these products. That sum = the expected utility of A.

It must be admitted that FP2, in virtue of that parenthetical remark in the second step, seems disturbingly inelegant, and lacks a certain precision that was an important feature of FP. It would be nice if we could come up with a variation on FP2 that gets around at least the second of these problems. But we have already seen that the correct solution to The Two Envelope Problem involves recognizing that we must revise FP, so that it says something about taking into account all relevant information. In light of this, I am inclined to think that the solution to The Two Envelope Problem is to replace FP with something along the lines of FP2.[8]

[8] Earlier versions of this paper were presented at the University of Massachusetts and The University of Rochester in 2005. I am grateful to members of both audiences, and to Ben Bradley, Stuart Brock, Kenny Easwaran, Adam Elga, Alan Hájek, Frances Howard-Snyder, Hud Hudson, Lee Markosian, Jonathan Schaffer, Ted Sider, Ryan Wasserman, and Catherine Wilson for helpful comments or discussion.

THE CASE FOR RATIONAL UNIQUENESS

Jonathan MATHESON

ABSTRACT: The Uniqueness Thesis, or rational uniqueness, claims that a body of evidence severely constrains one's doxastic options. In particular, it claims that for any body of evidence E and proposition P, E justifies at most one doxastic attitude toward P. In this paper I defend this formulation of the uniqueness thesis and examine the case for its truth. I begin by clarifying my formulation of the Uniqueness Thesis and examining its close relationship to evidentialism. I proceed to give some motivation for this strong epistemic claim and to defend it from several recent objections in the literature. In particular I look at objections to the Uniqueness Thesis coming from considerations of rational disagreement (can't reasonable people disagree?), the breadth of doxastic attitudes (can't what is justified by the evidence encompass more than one doxastic attitude?), borderline cases and caution (can't it be rational to be cautious and suspend judgment even when the evidence slightly supports belief?), vagueness (doesn't the vagueness of justification spell trouble for the Uniqueness Thesis?), and degrees of belief (doesn't a fine-grained doxastic picture present additional problems for the Uniqueness Thesis?).

KEYWORDS: disagreement, evidence, justification, rationality

1. Introduction

Suppose that you have a body of evidence. Given that body of evidence, precisely how open are your doxastic options, rationally speaking, regarding any given proposition? Does this body of evidence rationally permit you to believe whatever you want or are the doxastic restrictions much more stringent? The Uniqueness Thesis claims that one's doxastic options are rationally quite constrained by any given body of evidence. In fact, according to the Uniqueness Thesis there is at most one rational doxastic attitude to adopt towards any one proposition given any particular body of evidence.

The Uniqueness Thesis, or rational uniqueness,[1] is a strong claim about the permissiveness of rationality. Although it has been put to much use in the current

[1] I will be using the terms 'Uniqueness Thesis' and 'rational uniqueness' interchangeably. 'The Uniqueness Thesis' is Richard Feldman's term ("Reasonable Religious Disagreements," in *Philosophers without God: Meditations on Atheism and the Secular Life*, ed. Louise Antony (Oxford: Oxford University Press, 2007), 194-214), whereas David Christensen ("Epistemology of Disagreement: The Good News," *Philosophical Review* 116 (2007): 187-218) uses 'rational uniqueness.'

debates regarding the epistemology of disagreement, the Uniqueness Thesis tends to receive rather little explicit attention.[2] In what follows I will explain and clarify a version of the Uniqueness Thesis which I will be defending. Although this characterization of the Uniqueness Thesis will differ from other accounts in the literature, it is a more general characterization which adequately captures the claim of rational uniqueness. I will then examine the relationship that this claim has with evidentialism, and proceed to defend my characterization of the Uniqueness Thesis from several objections.

2. Some Clarifications

First, what is the Uniqueness Thesis? As mentioned above, the claim is that given a body of evidence, there is no more than one justified doxastic attitude to have toward a proposition. I will be defending a precisification of this thesis which claims the following:

> (UT) For any body of evidence E and proposition P, E justifies at most one doxastic attitude toward P.

Some clarifications are in order to make this claim precise. First, bodies of evidence are possessed by individuals at times. They can be shared by more than one individual, and they can change from time to time. (UT), however, makes no reference to individuals or times since (UT) claims (in part) that *who* possesses the body of evidence, as well as *when* it is possessed, makes no difference regarding *which* doxastic attitude is justified (if any) toward any particular proposition by that body of evidence.

(UT) concerns justification. In what follows, I will be using the terms 'justified,' 'reasonable,' and 'rational' all interchangeably unless otherwise noted. My concern will be with *epistemic* justification, so when I speak of a belief's reasonability or rationality it will concern the reasons to think that it is true, and not any means-ends or pragmatic considerations to which 'rationality' or 'reasonability' are sometimes used to refer. (UT) concerns propositional justification, rather than doxastic justification.[3] That is, the kind of justification

[2] Roger White, "Epistemic Permissiveness," *Philosophical Perspectives* 19 (2005): 445-459, and Nathan Ballantyne, E.J. Coffman, "Uniqueness, Evidence, and Rationality," *Philosophers Imprint* (forthcoming) are notable exceptions.

[3] Doxastic justification concerns not only the factors which support adopting a certain doxastic attitude toward a proposition, but also how those factors are utilized in the formation of that doxastic attitude. Thus, doxastic justification concerns the status of a doxastic attitude which is held by an individual and depends in part upon how that individual came to have that doxastic attitude or upon what that doxastic attitude is based.

relevant to (UT) is solely a relation between a body of evidence, a doxastic attitude, and a proposition. How individuals have come to have the doxastic attitudes they have toward the proposition in question will not be relevant to our discussion. Further, individuals can be propositionally justified in adopting attitudes toward propositions which they psychologically cannot adopt. An ability to take on a given doxastic attitude toward a proposition is not a requisite for being propositionally justified in believing it. It is easy to see that (UT) concerns propositional justification and not doxastic justification since no mention is made of any particular individual actually having any particular doxastic attitude toward any proposition.

Importantly, it is not a necessary condition for being justified in believing p that one be able to demonstrate that one is justified in believing p. The *project* of justifying or giving a defense of one's belief that p is distinct from the *state* of being justified in believing that p. One need not be able to articulate one's reasons for believing p in order to be justified in believing p.[4]

The version of the Uniqueness Thesis which I have given differs from other accounts in the literature in that it claims that *at most* one doxastic attitude is rational as opposed to claiming that there is *exactly one* doxastic attitude which is rational.[5] In most cases there will be exactly one rational doxastic attitude, but as Feldman[6] notes, it may be that no doxastic attitudes are rational to adopt towards a proposition which one does not or cannot understand. That is, it may be that a necessary condition for *any* rational doxastic attitude towards a proposition that one understands or grasps, or at least is able to understand or grasp, the proposition in question. So, to at least avoid taking a stand on that issue, I think that the Uniqueness Thesis is best stated in this slightly weaker way.

It is important to clarify how a body of evidence should be understood. (UT) makes a claim about how many doxastic attitudes are justified toward a proposition by a body of evidence. (UT) does not make any claims regarding how many doxastic attitudes are justified toward a proposition by distinct total bodies of evidence which contain identical bodies of evidence as proper parts. Both Earl Conee[7] and

---

[4] See James Pryor, "The Skeptic and the Dogmatist," *Nous* 34 (2000): 517-549, for a more detailed defense of this claim.

[5] Contrast White, "Epistemic Permissiveness," and Thomas Kelly, "Peer Disagreement and Higher Order Evidence," in *Disagreement*, eds. Richard Feldman and Ted Warfield (New York: Oxford University Press, 2010), 111-174, for two examples.

[6] Richard Feldman, "Epistemological Puzzles about Disagreement," in *Epistemology Futures*, ed. Stephen Hetherington (New York: Oxford University Press, 2006), 216–36.

[7] Earl Conee, "Rational Disagreement Defended," in *Disagreement*, eds. Richard Feldman and Ted Warfield, 69-90.

Alvin Goldman[8] have illustrated that there can be cases where individuals who differ in terms of their higher-order evidence (their evidence about the character of their first-order evidence) can be justified in adopting distinct doxastic attitudes toward a single proposition, even when these disparate bodies of evidence share the same first-order evidence (or 'evidence relevant to the dispute' as Conee refers to it). Conee and Goldman each give examples of cases where two individual's have the same first-order evidence regarding a proposition, but have different evidence regarding the various relevant epistemic principles (or 'E-systems' as Goldman talks of them). Given all of this, the two individuals are justified in adopting distinct doxastic attitudes toward the relevant proposition. These considerations, however, do not tell against (UT). (UT) is not a claim restricted to one's first-order evidence (or evidence directly pertaining to the dispute). (UT) claims that a body of evidence supports at most one doxastic attitude toward a proposition; it makes no claim whatsoever regarding the number of doxastic attitudes one could be justified in adopting towards that proposition if that single body of evidence where supplemented in distinct and diverging ways resulting in disparate total bodies of evidence.

Finally, what are the doxastic options which (UT) concerns? Often doxastic attitudes are seen as 'all-or-nothing' affairs and are limited to three possibilities: belief, disbelief, and suspension of judgment. Others think of the doxastic options in a more fine-grained way and speak instead of degrees of belief. Those within the degreed camp can be further distinguished by how expansive each doxastic attitude is conceived to be, from a single point value or probability function, to a range of probability functions. (UT) makes no claim regarding which doxastic picture is correct. Although (UT) is silent on this matter, in the bulk of the paper I will be examining the prospects of (UT) given a tripartite doxastic taxonomy, but I will conclude by briefly extending the discussion to a richer doxastic picture.

3. Uniqueness and Evidentialism

Before going any further it will be beneficial to briefly examine the relationship between (UT) and Evidentialism. Evidentialism is the claim that which doxastic attitude one is justified in adopting toward a proposition at a time is determined entirely by one's evidence at that time. More formally the evidentialist thesis is as follows:

[8] Alvin Goldman, "Epistemic Relativism and Reasonable Disagreement," in *Disagreement*, eds. Richard Feldman and Ted Warfield, 187-215.

> (ET) For any subject S, proposition P, time T, and doxastic attitude D, S is justified in adopting D toward P at T if and only if having D toward P fits the evidence S has at T.[9]

Evidentialism is thus a supervenience claim – it claims that which doxastic attitude is justified for an individual regarding a proposition at a time supervenes upon that individual's evidence at that time.

(ET) might be thought to entail (UT), but this is not the case. Even if which doxastic attitude is justified for an individual at a time is entirely determined by that individual's evidence at that time, it needn't be that there is no more than one competitor doxastic attitude which is so justified. That is, it could be that although evidence alone determines which doxastic attitude(s) are justified for me, I nevertheless have doxastic options.[10] It could be that my evidence justifies a set of options such as the disjunctive option belief or suspension of judgment, where either of these attitudes would be justified for me. It could be that what supervenes on the evidence is broader than any one doxastic attitude. So, it could be that a body of evidence is such that believing or suspending judgment would be justified for any individual with that body of evidence – that it is this disjunction of doxastic attitudes which supervenes on the evidence.[11] The evidential thesis does not rule this out. I will critically examine below whether a body of evidence could be like this, but for now my task is simply to illustrate that one could consistently endorse evidentialism and yet deny (UT).[12]

Similarly, (UT) might incorrectly be thought to entail (ET).[13] If one's evidence always picks out at most one rational doxastic attitude concerning a proposition, then it must be that the evidence (and the evidence alone) is what is conferring the justificatory status upon the doxastic attitude. However, (UT) is silent as to *how* this uniquely rational attitude is determined by the evidence. (UT) is clearly consistent with (ET) in that it may be the attitude which *best fits* the evidence which is the uniquely justified one, but (UT) can also be consistently

---

[9] This closely resembles the (EJ) principle Conee and Feldman define and defend as evidentialism. See Earl Conee, Richard Feldman, "Evidentialism," *Philosophical Studies* 48 (1985): 15-34.

[10] Goldman makes such a suggestion. See Goldman, "Epistemic Relativism and Reasonable Disagreement."

[11] Alternatively, on a degreed picture, one's evidence might justify a larger range of confidence than any one doxastic attitude (even if the attitudes themselves are ranges and not precise points). In sum, what the evidence justifies might be broader than what the doxastic options are.

[12] In addition, the evidentialist could endorse a relativist notion of 'fit,' in which case she could endorse (ET) and deny (UT). This route is explore in Ballantyne, Coffman, "Uniqueness, Evidence, and Rationality."

[13] Ballantyne, Coffman, "Uniqueness, Evidence, and Rationality."

defended along with 'anti-evidentialism' which claims that the uniquely justified attitude is the one which has the *least or worst fit* with the evidence.[14] (UT) can also be consistently defended with 'Tuesday evidentialism which claims that the uniquely justified attitude is the one best supported by one's evidence obtained on Tuesdays as well as other such 'evidential' theories of justification. Anti-evidentialism and such other 'evidential' theories are not very plausible, but they do show that the defender of (UT) need not defend (ET) on pains of consistency.

Nevertheless, the implausibility of anti-evidentialism shows that there is an intimate relation between (UT) and (ET). If (ET) is false, then it seems that (UT) will also be false since anti-evidentialism and the other alternative 'evidential' theories are clearly false. So, the falsity of evidentialism would spell trouble for rational uniqueness.[15] The non-evidentialist has no business endorsing (UT).[16] However, I will not be examining indirect attacks to (UT) via criticisms of evidentialism, though such critiques do affect the plausibility of (UT), given the implausibility of anti-evidentialism and the other relevant alternatives.[17] Rather, in what follows I will be assuming the truth of evidentialism, or (ET), and will proceed to assess the prospects for (UT) given that assumption regarding the nature of epistemic justification. There are enough challenges to (UT) worthy of our consideration within this restricted domain.[18]

---

[14] Thanks to Richard Feldman for both pointing this relation out to me and giving me the term 'anti-evidentialism.'

[15] Presumably, rival accounts of epistemic justification will have their own rival accounts of a uniqueness claim where the feature or features they see as relevant for epistemic justification will be held fixed and replace the body of evidence in (UT). Thus, for process reliabilism we might expect something like:

(UT') For cognitive belief forming process C, subject S, proposition P, and time T, C justifies at most one doxastic attitude for S toward P at T.

Or,

(UT'') For any degree of reliability D had by a belief forming process C, Subject S, proposition P, and time T, the D of C justifies at most one doxastic attitude for S toward P at T.

Now it seems that the reliabilist would reject (UT'), although I'm not sure what she would say about (UT''). However, I will not here explore rival accounts of the Uniqueness Thesis nor their plausibility within rival accounts of epistemic justification.

[16] Is this itself a problem for (UT)? I don't think so, since most of those who have been concerned with the truth of (UT) are already committed to some form of evidentialism. The interesting question is whether (UT) it true even given the truth of evidentialism.

[17] See Ballantyne, Coffman " Uniqueness, Evidence, and Rationality," for such an attack on (UT).

[18] If the reader is not convinced that evidentialism is correct, then the rest of this paper can be treated as defending a conditional claim: if evidentialism is true, then the Uniqueness Thesis is correct. For more on the relationship between the Uniqueness Thesis and Evidentialism see Ballantyne, Coffman "Uniqueness, Evidence, and Rationality."

4. Motivating Rational Uniqueness

Before turning to these challenges to (UT), it is worth motivating this epistemic claim. Why think that a body of evidence justifies at most one doxastic attitude toward any proposition? (UT) seems quite plausible on a tripartite doxastic taxonomy (that our doxastic options are belief, disbelief, and suspension of judgment) since there are also three ways that a body of evidence can be regarding a proposition. It seems that a body of evidence either on balance supports p, on balance supports not-p, or is on balance neutral between p and not-p. A body of evidence cannot on balance support both p and on balance support not-p, and a body of evidence cannot on balance support p and on balance be neutral between p and not-p.[19]

So, given our evidentialist assumption, there seems to be a unique doxastic option regarding p for each type of possible evidential situation regarding p. If the evidence supports p, then believing p is the uniquely justified doxastic attitude regarding p. If the evidence supports not-p, then disbelieving p is the uniquely justified doxastic attitude regarding p. If the evidence is neutral between p and not-p, then suspending judgment is the uniquely justified doxastic attitude regarding p. If the body of evidence is neutral between p and not-p it is not that believing or disbelieving p would both be rational. Suspension of judgment is always an option, and it is this doxastic attitude that would be rationally demanded in such an evidentially tied situation.[20] A body of evidence is like set of scales: there are only three ways the scales can be: favoring the left side, favoring the right side, or perfectly balanced.

So, there are three ways that a body of evidence can be, and, at least given a tripartite doxastic taxonomy, one unique doxastic response which fits each evidential possibility. So, consideration of the tripartite account of doxastic attitudes and ways that a body of evidence can be give reason to believe that rational uniqueness is correct.

---

[19] Similarly a body of evidence cannot on balance support not-p and on balance be neutral between p and not-p. This motivation for rational uniqueness parallels one given by White, "Epistemic Permissiveness."

[20] In this way, one's doxastic options are more expansive than one's practical options always are. Thus, William James (*The Will to Believe and Other Essays in Popular Philosophy* (New York: David McKay, 1911)) is mistaken in likening one's doxastic options to a practical choice such as whether to offer a marriage proposal. While not deciding whether to offer the proposal gives the same result as deciding to not offer the proposal, in the doxastic cases, suspension of judgment offers a third distinct alternative. This third option is unparalleled in some practical cases (those James calls 'forced' options). For more on this point see Richard Feldman, "Clifford's Principle and James's Options," *Social Epistemology* 20 (2006): 19-33.

In what follows I will examine various objections to the Uniqueness Thesis and respond to each. I will consider objections concerning disagreements, being cautious, the breadth of justified doxastic attitudes, and problems concerning vagueness.

5. Objection 1: Disagreement

The Uniqueness Thesis is not without its detractors. Gideon Rosen, for instance, takes it as an obvious fact that individuals can reasonably disagree even given a single body of evidence. He explains,

> When a jury or a court is divided in a difficult case, the mere fact of disagreement does not mean that someone is being unreasonable. Paleontologists disagree about what killed the dinosaurs. And while it is possible that most of the parties to this dispute are irrational, this need not be the case. To the contrary, it would appear to be a fact of epistemic life that a careful review of the evidence does not guarantee consensus, even among thoughtful and otherwise rational investigators.[21]

Here Rosen takes it that cases of disagreement show that rational uniqueness is clearly false. Cases of disagreement have been used in ethics to argue that there is no universal objective moral code, and it seems that Rosen takes it that a similar case can be made against rational uniqueness.

Rosen, however, thinks that there is an important asymmetry between the epistemic and ethical cases of disagreement. He thinks that even if ethical disagreement does not tell against ethical objectivism, the same is not true regarding the effect of disagreement upon rational uniqueness. He writes,

> Rational permission differs from moral permission in the following respect. There is no presumption that when an act is morally impermissible, we should be able to lead any clear-headed, open-minded, intelligent agent to see that it is. That's why rationally irresolvable moral disagreement is a possibility. In the epistemic case, on the other hand, a claim to the effect that one is obliged to follow a certain rule is undermined if we can describe a reasonable-seeming, fully reflective, and fully livable human practice that eschews it ... You can charge [the disagreeing party] with irrationality, and they will listen to the indictment. But what will you say to back it up? When they ask you, "What's wrong with our way of proceeding?", what will you say? If you have nothing to say, then the charge will not stick. Not only will they (quite reasonably) fail to heed you. If you have nothing to say, then in my view the charge is mistaken.[22]

[21] Gideon Rosen, "Nominalism, Naturalism, and Epistemic Relativism," *Philosophical Perspectives* 15 (2001): 71.

[22] Rosen, "Nominalism," 83.

There are a couple of things worth noting about Rosen's comments. First and foremost, Rosen seems to confuse the *project of justifying one's beliefs* with the *state of having justified beliefs* or being justified in believing certain things. As mentioned above, being epistemically justified in believing p does not entail that one is able to convince others that p or to even give a non-question begging defense of p. Being unable to articulate why one's claim is correct, or what is wrong with an opponent's counter-argument, does not have the consequence that one's claim is not justified or that the opponent's claim is not unjustified. Given (ET), whether either claim is justified will depend upon its fit with the evidence, so one's inability to convince another does not show that the belief in question is not justified; thus, it does not show that (UT) is false.[23]

Second, it is not clear that morality permits rationally irresolvable disagreements in the way Rosen imagines – John Rawls[24] and Roderick Firth[25] at least think otherwise. Rawls and Firth are concerned with rational procedures and thus are using 'rational' in a sense other than we are, but if we closely examine Rosen's comments, it seems that he too is thinking of rationality in this procedural way. In the first quote given above, Rosen refers to each individual 'carefully weighing the evidence,' and 'being thoughtful.' Such responsible inquiry and careful consideration, however, does not guarantee that one's resulting beliefs will be epistemically justified.[26] As we are understanding it, epistemic justification is a matter of evidential fit alone. So, even if ideal agents could undergo rational procedures and come to different conclusions, this is not to say that those distinct conclusions are each epistemically justified by their evidence. To maintain that this is the case would be contrary to our evidentialist assumption.

I agree with Rosen that individuals can each undergo *rational procedures* and come to different conclusions, but this is not the type of rational permissiveness that is relevant to (UT). So long as ideal agents are still fallible judges of the evidence, two such individuals could each carefully weigh the same body of evidence and yet come to distinct conclusions. This, however, does not show that the body of evidence does not support only one of the differing doxastic attitudes.

---

[23] See Christensen, "Epistemology of Disagreement," and Pryor " The Skeptic and the Dogmatist," for more on this claim.

[24] John Rawls, "Outline for a Decision Procedure in Ethics," *Philosophical Review* 60 (1951): 177-197, John Rawls, *A Theory of Justice* (Cambridge, Mass.: Harvard University Press, 1971).

[25] Roderick Firth, "Ethical Absolutism and the Ideal Observer Theory," *Philosophy and Phenomenological Research* 12 (1952): 317-345.

[26] A similar distinction is made by Russ Shafer-Landau and utilized in attacking an anti-realist argument from moral disagreement. See Russ Shafer-Landau, *Moral Realism: A Defense* (New York: Oxford University Press, 2003), especially Ch. 9.

So, it is hard to see how Rosen's considerations regarding disagreements show that (UT) is false. Importantly, (UT) does not claim that it will always be easy to determine which doxastic attitude is the uniquely justified one to adopt toward a proposition given a body of evidence. Often it will not. In fact, this difficulty in part explains the widespread and persistent disagreement which we often encounter in our lives over a variety of topics.

6. Objection 2: The Breadth of Justified Doxastic Attitudes

Alvin Goldman briefly questions the Uniqueness Thesis concerning potential mismatches between one's doxastic options and the prescriptions made by a body of evidence. Goldman claims that it seems unlikely that the correct epistemic principles will make doxastic prescriptions only in the narrowest doxastic categories (whichever these may be). Goldman notes that there are psychological limits as to how narrow or wide a doxastic state can be regardless of the correct doxastic taxonomy. Given that there is some narrowest doxastic state, Goldman questions why we should think that the correct epistemology will have it that solely one of these narrowest doxastic states will be justified by any body of evidence. Goldman thinks that this criticism applies equally well to a tripartite taxonomy of doxastic attitudes as to a degreed notion of belief. On the tripartite picture, Goldman maintains that the evidentially prescribed doxastic attitude could plausibly be a disjunctive category such as 'belief or suspension of judgment.'[27]

From considerations explored above, it is hard to see how Goldman could be correct on this point. On the tripartite doxastic taxonomy a body of evidence simply could not support one of Goldman's disjunctive doxastic categories. A body of evidence either supports a proposition, supports its denial, or is neutral on the matter. Given which way the body of evidence is, belief, disbelief, or suspension of judgment respectively will be the justified doxastic attitude to have toward that proposition. A body of evidence cannot both be neutral regarding p and also support p, so it is not plausible that Goldman's disjunctive categories could be justified for an individual to adopt toward a proposition given a body of evidence.

7. Objection 3: Borderline Cases and Caution

Thomas Kelly does not think that rational uniqueness is especially plausible even on a tripartite picture of one's doxastic options. He thinks that marginal cases cast

[27] Goldman, "Epistemic Relativism and Reasonable Disagreement." Goldman even offers invented labels for such states such as 'belension' and 'disbelension' – the disjunctive category of disbelief and suspension of judgment.

insurmountable doubt upon the claim that there is at most one rational doxastic attitude to adopt toward a proposition given a body of evidence.[28] Kelly asks us to imagine the following case:

> Suppose that the evidence available to me is just barely sufficient to justify my belief that it will rain tomorrow: if the evidence was even slightly weaker than it is, then I would be unjustified in thinking that it will rain. Suppose further that you have the same evidence but are slightly more cautious than I am, and so do not yet believe that it will rain tomorrow. It is not that you are dogmatically averse to concluding that it will rain; indeed, we can suppose that if the evidence for rain gets even slightly stronger, then you too will take up the relevant belief.[29]

Given this setup, Kelly thinks that it is by no means clear that the reader is being any less reasonable than Kelly in adopting her distinct doxastic attitude toward the proposition that it will rain tomorrow.

It is hard to follow Kelly here. After all, as Kelly claims, the evidence *is sufficient* to justify believing that it will rain. When one's evidence is like that, withholding belief is not the justified doxastic response – one has sufficient reason to believe the proposition at hand. Being cautious is often a good thing, but in this case exercising caution is causing an individual to adopt a doxastic attitude which is distinct from the doxastic attitude toward p justified by the reader's evidence. So, in this case, the reader has adopted the wrong doxastic attitude regarding the proposition that it will rain tomorrow given her evidence. This follows from our evidentialist assumption and Kelly's given description of the case.

This being so, nothing in (UT) dictates how bad of an epistemic crime it is to fail to adopt the uniquely correct doxastic attitude toward a proposition,

---

[28] Roderick Chisholm makes room for such cases. In setting out his terms of epistemic appraisal, Chisholm asserts that one such status is acceptable. He defines this epistemic status as follows: "h is acceptable for S =Df Witholding h is not more reasonable for S than accepting h." (Roderick Chisholm, *Theory of Knowledge*, 2nd ed. (Englewood Cliffs**,** NJ: Prentice-Hall, 1977), 9.) Chisholm goes on to claim that not all propositions that are acceptable, in this sense, are also beyond reasonable doubt (such that accepting them is more reasonable than withholding). The motivation that Chisholm gives for this claim comes from what he sees might be an adequate theory of perception. He states that such a theory of perception might require us to say, "if I have that experience which might naturally be expressed by saying that I 'seem to see' a certain state of affairs (e.g., 'I seem to see a man standing there'), then the state of affairs that I thus seem to perceive (the proposition that a man is standing there) is one that is , for me, ipso facto, acceptable. It may be, however, that although the proposition is thus acceptable, it is not beyond reasonable doubt; i.e., although withholding it is not more reasonable than believing it, believing it cannot be said to be more reasonable than withholding it." (Chisholm, *Theory of Knowledge* , 9-10.)

[29] Kelly, "Peer Disagreement," 10.

particularly in borderline cases. In borderline cases such as this one, Kelly's suspension of judgment seems to be much less of an epistemic transgression than one who remains skeptical even after much more positive evidence comes in on the matter. It would be a mistake, however, to confuse a slight failure of rationality with rationality. In the case described, Kelly can be properly reproached for not believing in accordance with the evidence, though he still does epistemically better than he might have. The defender of (UT) can satisfactorily account for this case.

8. Objection 4: Vagueness

These considerations of borderline cases of justification can lead us to think that justification and rationality are vague concepts. Individuals who endorse this claim would deny that there is a clear and distinct line separating when one is justified in believing a proposition and when one is not. The borders between cases where a doxastic attitude is justified and those where it is not may well be fuzzy and such that no amount of inquiry or conceptual analysis would settle whether that doxastic attitude toward the proposition in question was justified. In this way, the justification relation may be relevantly like the 'is close to' relation.

Why think justification is vague? First, it could be that it is vague whether a certain body of evidence supports a proposition or whether a given doxastic attitude fits that body of evidence. This seems to be particularly likely where there are a lot of evidential considerations at hand and there are defeaters and defeater-defeaters on both sides of the issue. In addition, it could also be vague whether a certain piece of evidence is part of a given body of evidence. Whatever the conditions for evidence possession are, it is quite plausible that at least some of them are vague.[30]

The vagueness of justification might be thought to create problems for (UT). If justification is vague, then there will be bodies of evidence such that it is not clear whether a certain doxastic attitude is justified or not regarding a proposition. Given that possibility, one might wonder how (UT) could be correct. One might be tempted to reason that if justification is a vague concept, then at least in these cases of indeterminacy, there can be distinct doxastic attitudes that are such that one would be justified in adopting either of them towards a proposition.

[30] Access internalists, for instance, claim that a necessary condition on something being part of one's evidence is that one is aware of it (or aware that it is a mental state which one is in). So, one's evidence is restricted to the mental states that one is aware of being in. However, it seems that it can be vague whether one is aware of being in a particular mental state. To see this we can examine a sorites series on awareness between clear cases of awareness and clear cases of unawareness. Although there are clear cases on both sides of the spectrum, there does not appear to be any sharp divide between cases of awareness and cases of unawareness.

Vagueness will not present a challenge to (UT) in cases where it is indeterminate whether a body of evidence justifies one doxastic attitude, yet it determinately does not justify all other doxastic attitudes towards that particular proposition. In such a case, it may be that the evidence justifies no doxastic attitude toward that proposition, but (UT) is consistent with this. Recall that (UT) claims only that *at most* one competitor doxastic attitude is ever justified by a body of evidence toward a proposition.

Things might be thought to be more problematic in cases where it is indeterminate whether a body of evidence justifies each of two distinct doxastic attitudes. However, such cases will only present a problem for (UT) if the indeterminacy has it that both of these doxastic attitudes are then justified for the relevant proposition by this body of evidence. However, it is hard to see why the indeterminacy of justification would have this result.

To give an analogous case, it can plausibly be vague whether a certain color patch is blue and at the same time vague whether that same color patch is green. This fact, however, does not suffice to show that the color patch is both blue and green at the same time. Similarly, it may be vague that a certain individual is tall and at the same time vague that the same individual is not tall. However, there is no obvious route from this fact to the individual's simultaneously being both tall and not tall. Some responses to vagueness deny bivalence, but the consequence examined here would be much more drastic than that. Such a response would seemingly permit the truth of contradictions! So, indeterminacy in cases of justification does not entail the permissiveness of rationality. Even if it is vague whether each of two distinct doxastic attitudes fits the evidence, it does not follow that each doxastic attitude is justified for that proposition given that body of evidence.[31] So, it is at best unclear how vagueness will present any challenge to (UT).

9. Degrees of Belief

Thus far we have been examining the prospects for rational uniqueness given a tripartite picture of one's doxastic options. These days, many epistemologists have opted instead for a more fine-grained doxastic picture, preferring to talk in terms of degrees of belief rather than all-out-belief or all-or-nothing-belief. It is worth briefly exploring the prospects of the Uniqueness Thesis given this richer doxastic picture.

The Uniqueness Thesis has been claimed to be implausible on such a richer doxastic picture,[32] but the reasons typically given are as those we have considered

[31] Christensen ("Epistemology of Disagreement," 192) makes a similar point.

[32] See Kelly, "Peer Disagreement."

above. While it might be more plausible to think that there are borderline cases and vagueness issues on this richer doxastic conception, what we have said above will apply equally well on this more fine-grained approach. So, I don't think that these are good reasons to abandon the Uniqueness Thesis on a degreed picture of belief either.

However, above we dismissed the 'breadth of doxastic attitudes' objection since a tripartite picture of one's doxastic option seemed to not let this objection off the ground. We saw that a body of evidence simply could not be such that it supported [belief or suspension of judgment] toward p. Nonetheless, on a degreed doxastic picture, it seems that this same response will not do. It may seem possible that a body of evidence not uniquely prescribe only one doxastic attitude when there are so many more distinct doxastic alternatives. Goldman doesn't find the tripartite account very convincing, so what about his worry applied to a richer doxastic picture?

In applying his objection to a degreed picture of doxastic attitudes, Goldman gives the example of an Intergovernmental Panel on Climate Change which made projections it deemed 'likely.' The panel's finding of the projections being 'likely' was to be interpreted as having a 66 – 90% chance of being correct. Such a claim seems to allow for quite a bit of leeway regarding a degree of confidence in the projections which would seemingly nonetheless be rational. Goldman does not explicitly claim that epistemic principles are similar in the relevant ways to such projections, but claims that if they are, then there is some doxastic permissiveness.

The problem here seemingly has to do with the fact that there may be several distinct ranges which all cover the degree(s) of belief justified by one's evidence. For instance, if we suppose that the justified doxastic attitude is a range, we might wonder whether any doxastic attitude which is a proper part of that justified range would itself be justified. For instance, if the evidence justified the range of .66 -- .9 belief that p, would competitor doxastic attitudes of believing to degree .74 and believing to degree .76 each be justified as well?

Similarly, competitor doxastic ranges that each encompassed the justified doxastic attitude might each be thought to be justified. So, if the evidence justified believing p to degree .75, one might think that having the doxastic attitude which ranges from .75 – .8 and the doxastic attitude which ranges from .7 – .75 would be equally rational. If so, then there would be more than one competitor doxastic attitude which one would be justified in adopting toward a proposition given a single body of evidence.

Are such distinct doxastic attitudes toward p justified by a single body of evidence? I don't think so. I am yet to be convinced that the existence of justified doxastic attitudes which are ranges of degrees of belief has such a consequence. If

one has adopted a doxastic attitude toward a proposition which is more expansive than what is justified by the evidence (as in the latter case), then that individual has not responded appropriately to the evidence. In such a case, one's doxastic attitude is too encompassing. There is a doxastic attitude which better fits the evidence. Similarly, if an individual has adopted a doxastic attitude toward p which is narrower than what the evidence supports (as in the former case), then that individual as well has not responded correctly to the evidence – a broader range of belief is called for. Here too there is a doxastic attitude which better fits the evidence. In both cases the individual in question could have had a doxastic attitude which better fits the evidence. It is *this* doxastic attitude (the one which best fits the evidence) which is the justified one to have toward the proposition in question. So, in Goldman's example, the doxastic attitude to have is the range of .66 – .9 belief that the panel's projections are correct. Any narrower or broader doxastic range would fail to appropriately fit the evidence.

Here again, the degree to which individuals whose doxastic attitudes fail to precisely line up with the justified doxastic range are believing irrationally will depend upon how much larger or smaller one's doxastic attitude is in comparison to the justified range of belief. It may be on this response that what is in fact justified for the individual is to believe the proposition question to the exact probability which his or her evidence supports the proposition in question (supposing there is some such probability). It may be that this is psychologically impossible, but it may nonetheless be what is epistemically required. As mentioned earlier, an individual having propositional justification for adopting an attitude toward a proposition does not entail that she is able to adopt that attitude toward that proposition. So, considerations of psychological limitations do not pose a problem for (UT). Again, more can be said about degrees of irrationality or unjustified doxastic attitudes that may be able to soften this blow. It needn't be that such individuals are to blame or that they ought to be censured. So, it is by no means clear that such a doxastic picture rules out the Uniqueness Thesis. We have seen no additional complications for rational uniqueness that come from adopting a more fine-grained doxastic picture.

10. Conclusion

We have seen some motivation for endorsing rational uniqueness and have fought off several objections to this claim. Rational uniqueness seems to be quite defensible given a tripartite doxastic taxonomy, and we have not seen any complications that arise from adopting a more fine-grained account of our doxastic options. So, the uniqueness thesis appears to be a strong, yet plausible claim regarding the permissiveness of rationality.

"LEBENDIGE FORMEN." ZU ERNST CASSIRERS KONZEPT DER "FORMWISSENSCHAFT"

Christian MÖCKEL

ABSTRACT: The following text analyses the essays Cassirer made in his published articles and in his manuscripts of the exile period in order to dismiss the possibility of a unique science based on the physical causality (Wiener Kreis) and thus to assert the autonomy and specificity of the *Kulturwissenschaft*. In the same time, it guards against the identification of cultural knowledge and historical knowledge (Windelband, Rickert). The specific way to build concepts within the *Kulturwissenschaften* targets the objective relation between the Particular and the General and is to be found in the consideration of the form or style based on the experience of the expression (*Ausdruckserleben*). History or biology offers fruitful analogies to the understanding of the concepts or methods of *Kulturwissenschaft.*

KEYWORDS: causal science, science of form, perception of the expression, law, form, structure

1. Vielfalt der Sinnordnungen und Weisen von Objektivität

Nachdem das philosophische Schaffen Cassirers in den ersten beiden Jahrzehnten des 20. Jahrhunderts dem Generalthema 'Funktionsbegriff statt Substanzbegriff' verpflichtet gewesen war, ist es in den 20er Jahren bekanntlich mit der Ausführung des innovativen Konzeptes einer 'Philosophie der symbolischen Formen' kulturellen Lebens befaßt. Neben den Untersuchungen zu einer Phänomenologie der Sprachform, der Form des mythischen Denkens und der Form naturwissenschaftlicher Erkenntnis werden in vielen Aufsätzen und Vorträgen zentrale Begriffe wie symbolische Form oder symbolische Prägnanz entwickelt. Das philosophische Grundproblem jener Zeit – das Verhältnis von Leben und Form –, das Cassirer fest im Blick hat, erfährt eine Lösung, die jegliche Antinomie vermeidet: "Für die Philosophie [...] kann [...] niemals das Leben selbst, vor und außerhalb aller Geformtheit, das Ziel und die Sehnsucht der Betrachtung bilden: sondern für sie bilden Leben und Form eine untrennbare Einheit."[1] Und dies bedeutet umgekehrt auch, daß keine Form selbst, vor und

[1] Vgl. Ernst Cassirer, "Der Begriff der symbolischen Form im Aufbau der Geisteswissenschaften," in ders. *Aufsätze und kleine Schriften. (1922-1926). Gesammelte Werke.* Bd. 16, Hrsg. von B. Recki (Hamburg: Meiner, 2003) [ECW 16], 75-104, 104.

außerhalb aller Lebendigkeit anzunehmen ist. "Diese Einheit [...] von Leben und Form macht [...] den eigentlichen Begriff des Geistes, sein 'Wesen' aus."[2] Die Philosophie der symbolischen Formen, die die "Urtatsache des [...] 'Lebens,'" das "Urphänomen des Lebens selbst" in "seinem Bestand und in seiner vollständigen Entfaltung" darzustellen sucht,[3] kann auch als eine Phänomenologie der lebendigen Formen bzw. des geformten Lebens entziffert werden. Seit 1925 ringt Cassirer darum, die Ausdrucksfunktion bzw. die Ausdruckswahrnehmung sozusagen als 'Schlußstein' seiner Philosophie einzufügen. Der Mensch als kulturschaffendes Lebewesen wird als dasjenige gefaßt, "was ausdrucksmäßig als 'Person' [...] bestimmt ist."[4] Damit scheint Cassirers Philosophie im Grunde ihre endgültige Form erlangt zu haben.

Doch in den außerordentlich produktiven 30er Jahren, vor allem in ihrer zweiten Hälfte, die er in der englischen und schwedischen Emigration verbringt, arbeitet Cassirer intensiv an einer Vertiefung, Neubegründung und Fortführung seiner Philosophie. An philosophisch Neuem sind neben den Versuchen über eine wechselseitige Begründung von symboltheoretischer Kulturphilosophie und philosophischer Anthropologie[5] vor allem seine Theorie der drei 'Basisphänomene' der Weltwahrnehmung und die sich daran anschließenden wissenschaftsphilosophischen Entwürfe zu nennen.[6] Die Wahrnehmung als die "Grund- und Urschicht aller Bewußtseinsphänomene"[7] werde konstituiert durch ein monadisches, ein personifizierend-verstehend-bewirkendes und ein versachlichend-kontemplatives Basisphänomen. Im Zentrum der Überlegungen zur Wissenschaft steht die Theorie einer eigenen Logik (d.h. Weise von Objektivität) der Kulturwissenschaften, die Cassirer in den miteinander verzahnten Konzepten von ursprünglicher Ausdrucksfunktion und 'Formwissenschaft' fundiert glaubt. Die Begründung dieser Grundunterscheidung zweier Wissenschaftslogiken wird im

[2] Vgl. Ernst Cassirer, *Zur Metaphysik der symbolischen Formen,* in ders. *Nachgelassene Manuskripte und Texte*. Bd. 1, Hrsg. von J. M. Krois (Hamburg: Meiner, 1995) [ECN 1], 267.

[3] Vgl. Cassirer, *Zur Metaphysik,* 263.

[4] Vgl. Ernst Cassirer, *Geschichte. Mythos*, in ders. *Nachgelassene Manuskripte und Texte*. Bd. 3, Hrsg. von K. Ch. Köhnke u.a. (Hamburg: Meiner, 2002) [ECN 3], 196.

[5] Vgl. Christian Möckel, "Kulturelle Existenz und anthropologische Konstanten. Zur philosophischen Anthropologie Ernst Cassirers," in *Zeitschrift für Kulturphilosophie* (Hamburg), 3. Jg. Heft 2 (2009): 209-220.

[6] Vgl. Christian Möckel, "Die Kulturwissenschaften und ihr 'Lebensgrund'. Zu Ernst Cassirers Beitrag zur Theorie der Kulturwissenschaften," in *Lebendige Form. Zur Metaphysik des Symbolischen in Ernst Cassirers 'Nachgelassenen Manuskripten und Texten'.* (Cassirer-Forschungen Bd. 13), Hrsg. von R. L. Fetz und S. Ullrich (Hamburg: Meiner, 2008), 179-195.

[7] Vgl. Ernst Cassirer, "Zur Logik der Kulturwissenschaften," in ders. *Aufsätze und kleine Schriften (1941-1946). Gesammelte Werke.* Bd. 24, Hrsg. von B. Recki (Hamburg: Meiner, 2007) [ECW 24], 357-490, 395.

phänomenalen 'Bestand' der Wahrnehmung aufgewiesen: in ihrer personifizierend-ausdruckverstehenden und in ihrer versachlichenden Auffassungs- bzw. Intentionalitätsrichtung. Beide Richtungen sieht Cassirer immer weiter auseinandertreten und sich verselbständigen. Das 'Urphänomen' der Versachlichung (*drittes* Basisphänomen) gilt ihm als Strukturmerkmal der Kausalwissenschaft und ihrer Dingauffassung, das 'Urphänomen' des Ausdrucksverstehens (*zweites* Basisphänomen) als Strukturmerkmal der Kulturwissenschaft und ihrer Werkauffassung. Soweit sich deren Logik als auf den Formbegriff orientiert erweist, greift Cassirer allerdings argumentativ hier ebenfalls auf das *dritte* Basisphänomen als einer objektiven 'Werk'-Konstitution zurück. Die Theorie der Basisphänomene erscheint damit ebensowenig aus einem Guß zu sein wie die auf ihr ruhende Wissenschaftsphilosophie. Der Begriff der Formwissenschaft findet außerdem methodisch Anwendung auf die Historie und auf die Biologie als deskriptiver Naturwissenschaft. Anschließend soll ein resümierender Blick auf den Begriff der Form als 'lebendiger Form' geworfen werden.

Cassirers wissenschaftsphilosophische Position läßt sich auf den Punkt gebracht wie folgt umschreiben: er zweifelt niemals an der Bedeutung und der Richtigkeit des Kausalprinzips für den Typus der theoretischen Naturwissenschaften als Gesetzeswissenschaften, polemisiert aber gegen die Reduktion von Wissenschaftlichkeit und Objektivität auf diesen *einen* Typus. Dies wird klar ausgesprochen, als er 1936/37 einen Beitrag über sein Verhältnis zur Philosophie des 'Wiener Kreises' vorbereitet. Cassirer positioniert sich in diesen Überlegungen gegen die vom Wiener Kreis favorisierte Einheitswissenschaft, die einzig und allein die physikalische Weise der Objektivität anerkennt und jegliches Wirklichkeitserleben auf diese zurückzuführen bestrebt ist. Demgegenüber besteht er darauf, daß "der 'Sinn' irgendwelcher Aussagen" durch eine Vielzahl von "System[en] von Kategorien bestimmt wird." Fehlender physikalischer Sinn bedeute keineswegs Sinnlosigkeit der Aussagen 'an sich.'[8] Neben dem "mathematisch-physikalischen Kausalbegriff" gebe es auch eine eigenständige biologische, historische und kulturwissenschaftliche Erkenntnis, was jeweils eigene Sinnordnungen impliziert.[9] Dennoch hält Cassirer grundsätzlich an der Vorstellung einer Einheit aller Wissenschaften fest, die in Gemeinsamkeiten wie dem Objektivitätsanspruch, der Zuordnung von Besonderem und Allgemeinem oder der Verschränkung von Kausal- und Formproblemen zum Ausdruck kommen.[10]

---

[8] Vgl. Cassirer, *Geschichte*, 196.

[9] Vgl. Cassirer, *Geschichte*, 241.

[10] Vgl. Christian Möckel, "Moritz Schlick und Ernst Cassirers Auseinandersetzung mit dem 'Wiener Kreis'," in *Moritz Schlick - Ursprünge und Entwicklungen seines Denkens. Schlickiana* Bd. 5, Hrsg von F.O. Engler und M. Iven (Berlin: Parerga, 2010), 207-224.

Vor diesem Hintergrund ist es nachzuvollziehen, daß Cassirer in den 30er Jahren wissenschaftsphilosophisch sozusagen an zwei Fronten kämpft. So verteidigt er 1936 in seinem Werk *Determinismus und Indeterminismus* das Gelten des Kausalitätsprinzips auch in der modernen Physik (Quantenmechanik), aus deren Entwicklung man einen physikalischen Indeterminismus ableitete, der das Kausalproblem habe obsolet werden lassen. Er setzt dem entgegen, daß die jeweilige Form der Kausalität und der jeweilige naturwissenschaftliche Objektbegriff in Wechselwirkung stünden und sich miteinander ändern können.[11] So ordne die moderne funktionalistische Betrachtungsweise den "Gesetzesbegriff [...] dem Gegenstandsbegriff vor [...], während er ihm früher nachgeordnet und untergeordnet war."[12] Außerdem sei durch die Quantentheorie das früher unlösbar scheinende Band zwischen "dem Kausalbegriff und dem Kontinuitätsprinzip" zwar aufgelöst worden, wodurch aber das Kausalprinzip nicht aufgehoben werde.[13] Auch unter den Bedingungen der modernen Quantentheorie gelte weiter, daß physikalische Urteile/Aussagen zunächst als individuelle 'Maßaussagen' auftreten, aus denen sich dann allgemeine 'Gesetzesaussagen' (Quantengesetz) ableiten, die schließlich zu universellen 'Prinzipien' als einem neuen Typus physikalischer Erkenntnisse (Quantenprinzip) vereinigt werden.[14] An der Spitze dieses Stufenmodells stehe – weiterhin – das 'Prinzip der Kausalität' als einer neuen methodischen Einsicht, die inhaltlich über das bereits Gewonnene nicht hinausgeht.[15] Niels Bohr habe in Wirklichkeit nicht "die Gültigkeit des Kausalbegriffs als solchem" bestritten, sondern nur die Möglichkeit, weiterhin "die Kausalbeschreibung der Phänomene direkt an die raumzeitliche Beschreibung zu binden." Ähnliches gelte letztlich auch für Heisenberg.[16] Interessant ist, daß Cassirer, wenn er vom System der physikalischen Erkenntnis und seiner Struktur spricht, anmahnt, diese Struktur unbedingt als "beweglich zu denken," als auf 'Plastizität' und 'Bildsamkeit' beruhend. Außerdem sieht er in seinem Werk über die physikalische Kausalität in der physikalischen Theorie eine "lebendige Form."[17]

Die zweite philosophische Kampflinie in den 30er Jahren bringt das Konzept 'Formwissenschaft' gegen die Alleinvertretungsansprüche der 'Gesetzeswissenschaft'

[11] Vgl. Ernst Cassirer, *Determinismus und Indeterminismus in der modernen Physik. Historische und systematische Studien zum Kausalproblem. Gesammelte Werke.* Bd. 24, Hrsg. von B. Recki (Hamburg: Meiner, 2004) [ECW 19], 13.
[12] Vgl. Cassirer, *Determinismus*, 159.
[13] Vgl. Cassirer, *Determinismus*, 193, 195.
[14] Vgl. Cassirer, *Determinismus*, 40, 49, 55, 59.
[15] Vgl. Cassirer, *Determinismus*, 69, 74.
[16] Vgl. Cassirer, *Determinismus*, 139, 147ff.
[17] Vgl. Cassirer, *Determinismus*, 91.

(Kausalität) ebenso in Stellung, wie gegen die Fehldeutung aller Wissenschaften jenseits der Naturwissenschaft als ausschließlich individualisierend-historischer oder axiologischer Erkenntnis. Dabei wird der Begriff der Formwissenschaft bei Cassirer sowohl auf das historische Leben und die Historie als Wissenschaft als auch auf das kulturell-geistige Leben und die sich etablierenden Kulturwissenschaften angewandt. Die methodologischen Überlegungen beider Richtungen nehmen auch das vegetative Leben und damit die Biologie als Formwissenschaft in den Blick. Sie gilt Cassirer nicht nur als Analogon beider Betrachtungsweisen, sondern zudem als Prototyp einer Verschränkung von formwissenschaftlicher und kausalwissenschaftlicher Betrachtungsweise, worauf am Ende noch einmal eingegangen wird.

Die Assoziation von Zusammengehörigkeit, die sich bei den Termini 'Philosophie der symbolischen Formen' und 'Formwissenschaft' einstellt, ist keineswegs zufällig oder unbeabsichtigt. Vielmehr zeigt sich gerade an diesem Typus von Wissenschaft, daß sie ohne die ideellen Formen der Philosophie orientierungslos bliebe, während die Philosophie in dem empirischen Material der Formwissenschaften wiederum ihr Material für die Formenschau findet. Und diese Beschäftigung mit den ideellen Formen spielt sich in drei großen Koordinatensystemen ab, die gemeinsame Schnittmengen besitzen oder die sich als unterschiedliche Dimensionen des Formlebens erweisen: in denen der Natur, der Kultur und der Geschichte. Gelegentlich spricht Cassirer auch vom sozialen Leben als einem eigenen Formensystem.

2. Der Formbegriff im Konzept der Formwissenschaften

Der Begriff der Form, der zweifellos *die* zentrale Kategorie in Cassirers gesamter Philosophie ist,[18] versteht sie sich doch als Formenlehre, als 'Morphologie' der Kultur, wird – trotz mancher Anläufe – nirgends systematisch aufgebaut und entwickelt. Sicher ist nur, daß er ihn in das die ganze Philosophiegeschichte durchziehende Formproblem gestellt wissen will. Welche Rückschlüsse erlaubt sein Konzept der Formwissenschaft auf diesen Schlüsselbegriff? Als Vorgriff und Einstimmung auf die nachfolgenden Ausführungen lassen sich fünf Aspekte nennen.

1. Der Formbegriff, der uns in Cassirers wissenschaftsphilosophischen Überlegungen entgegen tritt, besitzt eine holistische Dimension und steht für Struktur bzw. System, welche als Ganzheiten Charakter und Bedeutung ihrer

[18] "Wenn es überhaupt einen Begriff gibt, in dem man eine Gesamtsicht der Cassirerschen Philosophie fassen will, dann ist es der Begriff der Form – und zwar in seinem dynamischen Sinne als 'Werden zur Form', als Gestalten zur Form oder als das sich selbst eine Form Geben. Formwerden und Formgebung sind der Mittelpunkt des Cassirerschen Denkens." – Vgl. Oswald Schwemmer, *Ernst Cassirer. Ein Philosoph der europäischen Moderne* (Berlin: Akademie, 1997), 122.

Elemente prägen. Dies nennt Cassirer an anderer Stelle ebenfalls 'Prägnanz' der Form. Diesen eigenständigen und eigentümlichen Strukturen ist jeweils ein "inneres Gefüge"[19] bzw. ein 'Strukturgesetz' immanent, jeder Form ein bestimmtes Formprinzip, das die Richtung ihrer Wirkung, ihrer Darstellung, Repräsentation und Symbolisierung im und durch das konkrete Werk der Kultur, in und durch die konkrete biologische – oder historische – 'Lebensform' vorgibt. Da Richtung eine bestimmte 'Sinngebung' meint, in der und von der alle ihr zugeordneten Einzelphänomene ihren Sinn erfahren, und dabei gleichzeitig das Sinnganze repräsentieren, wäre Sinnordnung ein weiteres Synonym für Form. In diesem Zusammenhang qualifiziert Cassirer die Formen der Kultur auch als Weisen der "Motivierung" im symbolischen Tun des Menschen und als eigentümliche "Organe," eine Welt zu "sehen" und zu "bilden."[20] Das jeweilige Strukturgesetz der Form "färbt" auch auf Kategorien "ab," die - wie die des Raumes - in mehreren Sinnordnungen ihre Funktion entfalten und dabei unterschiedliche Bedeutungen (Synthesefunktionen) erlangen.[21] Den Grundformen der Kultur korrelieren nach Cassirer zudem "Funktionen" des menschlichen Geistes, seine "Energien," die sich in den Formen verwirklichen, entfalten, Sinnwelten, Sinnstrukturen "aufbauen." Cassirer wird sich in einem Vortrag 1945 über die moderne Linguistik zum Strukturalismus als "the expression of a general tendency of thought" bekennen.[22]

2. In ontologischem Sinne sind Formen ideelle Gebilde und führen keine eigene, gesonderte Existenz. Als Formen der Kultur oder Geschichte müssen sie vom Menschen, damit sie ihm während seiner bildenden Tätigkeit zur Verfügungen stehen können, 'entdeckt' oder 'erfunden' werden, müssen sich in seinen Werken und Lebensverhältnissen 'verkörpern,' 'objektivieren,' um danach als bleibende nicht mehr aus der Welt zu verschwinden.[23] Dieses 'Erfinden' der Formen meint kein bewußt-willentliches Hervorbringen, "wie es bei Taten oder auch bei Normen und Institutionen der Fall ist."[24] Sie sind für den Philosophen und Kulturwissenschaftler auch nur in diesen Verkörperungen aufzufinden und zu beschreiben, um einer reinen Formen-Analyse unterzogen werden zu können. In Geschichte und Kultur finden wir Zusammenhänge von reiner Form und Form in historischer, nationaler etc. Konkretion, Besonderung vor. Gleichzeitig haben wir

[19] Vgl. Cassirer, *Geschichte*, 221.
[20] Vgl. Cassirer, *Geschichte*, 188, 221.
[21] Vgl. Cassirer, *Geschichte*, 200.
[22] Vgl. Ernst Cassirer, "Structuralism in Modern Linguistics," in ders. *Aufsätze und kleine Schriften (1941-1946). Gesammelte Werke.* Bd. 24, Hrsg. von B. Recki (Hamburg: Meiner, 2007) [ECW 24], 299-320, hier 320.
[23] Vgl. Cassirer, *Geschichte*, 206.
[24] Vgl. Cassirer, *Geschichte*, 220.

es mit einer Stufung oder Vielzahl von Formen zu tun. Auch bilden die Formen selbst ein Ensemble, eine Totalität, eine Ordnung, die sowohl als reine wie auch als in historischer etc. Besonderung zu erfassen ist. In Natur, Gesellschaft und Kultur umfaßt eine konkrete 'Lebensordnung' eine Vielzahl einander korrelierender 'Lebensformen.' Formen sind nicht einzeln, isoliert voneinander zu erkennen: Formen "reflektieren sich ineinander" und werden "durch einander erkannt."[25]

3. Die Form eines kulturellen, historischen oder natürlichen Sinngefüges ist bei Cassirer grundsätzlich eine 'lebendige Form.' Was bedeutet diese immer wieder verwendete Auszeichnung der Form durch ihre Lebendigkeit, ist sie doch kein Lebewesen? Um das Überwinden der organischen Schranken des Menschen durch und in den Formen, in die er sein "Wirken [...] verwandelt," zu betonen, nennt sie Cassirer sogar 'anorganische' Formen.[26] Offenbar bedeutet 'lebendig' ein Mehrfaches, mindestens vier Sinne drängen sich aus den ausgewerteten Texten auf. Eine Bedeutung meint die Spontaneität der Form, ihr aktives Gestalten und Organisieren des jeweiliges 'Stoffes,' Gehaltes.[27] Ein weiteres, zweites Moment von lebendiger Form sieht Cassirer in ihrem unauflösbaren Bezug zur Ausdrucksfunktion: alles Lebendige drückt sich in Geformtem aus und das, was wir verstehen, verstehen wir als das geformt Ausgedrückte eines anderen Lebens.

Eine dritte, sehr wichtige Bedeutung, steht für das paradoxe Verhältnis von Bestand, Beharrendem, 'Starrheit,' 'Ewigkeit' (Identität) der Form und ihrer 'Beweglichkeit,' Veränderlichkeit, Wandlungsfähigkeit. Die "gewissen, relativ-gleichbleibenden Formen" gebären sich in immerzu "neuer Gestalt."[28] Bei der Wandlungsfähigkeit unterscheidet Cassirer organisches Werden, d.h. bloßen "organischen Wandel" im "bloß-organischen Sein" vom geschichtlichen Werden, d.h. von "Bedeutungs-Wandel" in der Kultur.[29] Damit plädiert er für einen "'dynamischen' Formbegriff," wonach die Formen "in keinem Moment sich selbst gleich[en]," sondern 'flüssig' werden, "ohne in dieser Flüssigkeit zu 'verschwimmen.'"[30] Diese Eigenschaft charakterisiert sie als "bleibende 'Gestalten'" mit "'lebendigen'" Zügen, "als werdend und wachsend, sich bildend und umbildend," als ihr Sein in der Geschichte habend. Das in jeder Form Beharrende nennt Cassirer auch ihr Moment der "Permanenz,"[31]

[25] Vgl. Cassirer, *Geschichte*, 67.
[26] Vgl. Cassirer, *Geschichte*, 206f.
[27] Vgl. Cassirer, *Determinismus*, 91.
[28] Vgl. Cassirer, *Geschichte*, 74, 207ff.
[29] Vgl. Cassirer, *Geschichte*, 208.
[30] Vgl. Cassirer, *Geschichte*, 223f.
[31] Vgl. Ernst Cassirer, *Kulturphilosophie. Vorlesungen und Vorträge. 1929-1941*, in ders. *Nachgelassene Manuskripte und Texte*. Bd. 5, Hrsg. von R. Kramme (Hamburg: Meiner, 2004) [ECN 5], 127.

das aber ihrer Wandelbarkeit keine endgültige Grenze setze. Die kulturellen Formen befinden sich durch das in ihnen Schaffen und Bilden der Individuen, was sie permanent wandelt und umbildet, in einem beständigen "Werden zur Form."[32]

Und schließlich meint lebendige Form viertens den von Cassirer – im Anschluß an Goethe und Hegel – vorausgesetzten Tatbestand einer ideellen Wandlung, Genesis, die systemimmanente Stufen durchläuft, was gleichzeitig einen Formwandel (Metamorphose) bedeutet. Überhaupt denkt Cassirer in strukturellen Stufenmodellen,[33] die er nicht zuletzt Goethe und Hegel entlehnt: so in den Strukturen Ausdruck-Darstellung-Bedeutung; Wahrnehmung-Anschauung-Denken; Mythos-Sprache-Wissenschaft, aber auch Nachahmung-Manier-Stil. etc. Das Strukturgesetz bzw. das Formprinzip regulieren dabei den inneren Formwandel und den äußeren Gestaltwandel. Alle Entwicklung einer Form könne nur *in* dem Kreise ihres ideellen Seins aufgewiesen werden, nicht aber die Entstehung dieses Kreises *aus* einem anderen Kreis des Seins.

5. Es fällt außerdem auf, daß für Cassirer jegliche analytische Erforschung der Formen durch reine Kontemplation erfolgen muß und keineswegs auf Wollen und Wirken zielen darf. Die "reine Kontemplation der 'Formenfülle'" habe sich von allen Aspekten der 'bloßen Aktion' (Wollen/Willen und Wirkung) fernzuhalten, aber eben auch von theoretischer Erklärung. In die "Welt der Formen" müssen wir uns "schauend versenken, um ihren Reichtum, ihre Fülle, ihren Sinn zu verstehen."[34] Selbst in Sprach- und Kunstgeschichte betreibe der Form-Historiker vor allem kontemplative "Sichtbarmachung vergangener Formwelten." Anderenorts nennt Cassirer diese reine Kontemplation die 'Wendung zur Idee', die "die Synthese von Theorie und Praxis" enthalte.[35]

3. Form, Formanalyse, Formwissenschaft

a. Geschichtsbetrachtung

Wenden wir uns zunächst den nachgelassenen Entwürfen aus den 30er Jahren (ECN 3) zu, die die eigenständige Weise der Sinn-Betrachtung durch die 'historische Erkenntnis' aufdecken sollen. Wenn Cassirer darauf verweist, daß das Eigentümliche der historischen Erkenntnis nicht in einer besonderen "logical

[32] Vgl. Cassirer, *Kulturphilosophie,* 192.
[33] Vgl. Cassirer, *Zur Metaphysik*, 262.
[34] Vgl. Cassirer, *Geschichte,* 219, 221.
[35] Vgl. Cassirer, *Zur Metaphysik*, 191.

structure of historical thought" zu suchen sei,[36] dann richtet sich dies gegen die Versuche, der 'Logik der naturwissenschaftlichen Begriffbildung' eine eigene, nicht-reduzierbare "special logic of history" (Windelband, Rickert) gegenüberzustellen.[37] Diesen Platz dürfe allein eine 'Logik der kulturwissenschaftlichen Begriffsbildung' besetzen. Alle übrigen Typen der Wissenschaft setzen sich aus unterschiedlich proportionierten Relationen beider Betrachtungsweisen zusammen. Die Mss. der 30er Jahre verfolgen einen breiten, systematischen Ansatz. Danach rekonstruiert und versteht der Geschichtsschreiber das 'historische Leben,' das sich in Monumenten und Dokumenten niederschlägt, in den drei Dimensionen der Weltwahrnehmung, die die Konstitutionsrichtungen der drei Ur- oder Basisphänomene realisieren: Dabei kommen Auto-Biographie, politische Geschichte und Werk- bzw. Formgeschichte heraus.[38] Auf die ersten beiden Weisen, Geschichte zu schreiben, kann und soll hier nicht weiter eingegangen werden. Ich beschränke mich auf die dritte Weise, die darauf fußt, daß die urphänomenale, nicht reduzierbare und nicht ableitbare Werk-Wahrnehmungsrichtung eine Welt der objektiven Formen eröffnet, wodurch ein Werk der Kultur als einer Sinnordnung bzw. einer bestimmten Stilrichtung etc. zugehörig entschlüsselt werden kann.

Dem Historiker stellt sich hier die Aufgabe, Werk-Geschichte als Form-Geschichte zu betreiben. Dies bedeute, sowohl die einzelnen historischen Monumente, die eine 'ewige' Form auf individuelle Weise ausdrücken, zu enträtseln, als auch die Geschichte ganzer Kultur- und Sinnformen zu beschreiben und zu deuten, also Sprachgeschichte, Kunstgeschichte etc. zu schreiben. Cassirer charakterisiert die Formen, mit deren Geschichte bzw. mit deren einzelnen Verkörperungen es der Historiker – auf individuelle Weise – zu tun hat, als Ganzheiten, Gestalten und Strukturen. Damit der Historiker 'bestimmte Formwelten,' also die Sphären der Kultur und ihre Strukturen 'entdecken' kann, muß er sich der reinen Form-Analyse des Kulturwissenschaftlers und des Philosophen bedienen bzw. versichern. Die vergangenen "'wirklichen' Lebensformen" erfahren so ihren Sinn, ihre Verstehbarkeit aus der jeweiligen Formwelt, in die sie gehören. Zu den Formwelten gelangt man durch die verschiedenen "Organe des Sehens," die jeweils ein anderes "Universum" der Kultur erstehen lassen.[39] Diese Vielfalt des sinnerschließenden 'Sehens' vollzieht der Historiker auch innerhalb einer Form oder Formwelt, so der der Kunst, was ihn eine Vielzahl von Kunststilen entdecken

[36] Vgl. Ernst Cassirer, *An Essay on Man. An Introduction to a Philosophy of Human Culture* (New Haven: Yale Press, 1947), 177.
[37] Vgl. Cassirer, *An Essay,* 186.
[38] Vgl. Cassirer, *Geschichte,* 3, 14ff.
[39] Vgl. Cassirer, *Geschichte,* 31.

und verfolgen läßt. Die Werk-Geschichte als "Kulturgeschichte" hat das "Eigenleben," die "eigene Entwicklung" jeder dieser Formen zu beachten und mit darzustellen.[40]

Gleichzeitig habe sie die "Formbeziehungen" in den Blick zu nehmen. Dies kann sie auf zweierlei Weise tun: einmal mit Hilfe der der Marxschen und Weberschen Soziologie entstammenden, dabei nicht kausal aufgefaßten "Formenlehre der Geschichte" ("historische Formenlehre"), die insbesondere den "Korrelationen zwischen den Formen der Wirtschaft einerseits und den Formen des Rechts, des Staates - auch der Kunst und Wissenschaft" - nachspürt.[41] Die soziologische Formenlehre habe zudem zum Thema, wie die korrelativen Formen "jeweilig in einer bestimmten gesellschaftlichen und kulturellen Ordnung zusammengefaßt sind." Eine solche Ordnung umschreibt Cassirer gelegentlich mit dem Dilthey entlehnten Terminus 'Lebensordnung'. Bei dieser Art Betrachtung des historischen Formengefüges habe die "kausale Frage, welches [in ihm - C.M.] das eigentlich 'begründende' Glied" sei, ganz zurückzutreten, da die "Formen [...] sich in einander [reflektieren] und [...] durch einander erkannt" werden.[42] Diese aus der soziologischen "Strukturlehre der Gesellschaft" hervorgehende Formlehre könne nicht zuletzt deshalb "als *eine* der 'Prinzipienwissenschaften' der Geschichte" angesetzt werden, weil auch die Historie im historischen Ablauf der Phänomene "gewisse, relativ-gleichbleibende Formen erkennen" wolle.[43]

Aber auch die Philosophie der symbolischen Formen ist nicht nur auf die statische Analyse einzelner Formen sondern zugleich "auf das Ineinandergreifen der Totalität der Formen gerichtet,"[44] was dem Historiker Formbeziehungen an die Hand gibt. Darüber hinaus sieht Cassirer zwischen der historischen Erkenntnis und der analytischen Erkenntnis des Strukturgesetzes der einzelnen symbolischen Formen ein 'doppelseitiges Verhältnis' bestehen. Einerseits kläre das historische Denken, indem es die symbolischen Formen "in ihrem Werden betrachtet," den Philosophen darüber auf, was jede dieser Formen "ist," erschließe es "ihre Wesensart, ihre spezifische Eigentümlichkeit." Müsse doch die "rein statische [Form-] Analyse [...] immer durch eine genetische [Betrachtung - C.M.] ergänzt werden," die allerdings "die Wesens-Betrachtung (die Form-Analyse) nicht ersetzen und verdrängen" kann.[45] Damit verhindere die historische Betrachtung, daß die Strukturerkenntnis ausschließlich als "Betrachtung 'toter' (statischer) Formen"

[40] Vgl. Cassirer, *Geschichte,* 58.
[41] Vgl. Cassirer, *Geschichte,* 67.
[42] Vgl. Cassirer, *Geschichte,* 67.
[43] Vgl. Cassirer, *Geschichte,* 74.
[44] Vgl. Cassirer, *Geschichte,* 183.
[45] Vgl. Cassirer, *Geschichte,* 175.

vonstatten geht. Vielmehr führe sie bis zu dem Punkt, wo wir in ihr das Leben der jeweiligen Form erfassen. Andererseits sind die

> einzelnen symbolischen Formen, wenn wir sie [...] in ihrer Ganzheit, die 'Bestand' und 'Veränderung' zugleich in sich schließt, erfassen, nun auch wieder die Organe, kraft deren uns historisches Leben in seinem ganzen Umfang und in seinen bewegenden Kräften, in seinen Ur-Motiven erst zugänglich wird[.] Jede dieser Formen führt uns [...] zu einer Urschicht menschlichen Denkens, Fühlens, Wollens zurück[,] von wo aus wir das Ganze derjenigen menschlichen Bewegungen, Handlungen, Taten, die wir mit dem Namen 'Geschichte' bezeichnen, erst ganz verstehen können.[46]

Damit erschließt die Philosophie die eigentliche "Motivierung menschlichen Handelns," was wiederum den Gang der Geschichte auf "bestimmte universelle Strukturgesetze" beziehen läßt.[47] Form-Analyse und Form-Geschichte greifen also ineinander.

Das Verhältnis der Form-Analyse zur historischen Erkenntnis der Form sieht Cassirer folglich darin, "daß a) jede 'Form' ihre Geschichte hat und daß die Gesamtheit der Phänomene, in denen sie sich konkret darstellt, uns nicht anders denn im 'Werden,' in geschichtlicher Entwicklung gegeben ist."[48] Und diese bedarf auch der Kausalbetrachtung. Gleichzeitig operiert b) die "rein statische [philosophische – C.M.] Form-Analyse" "mit ganz anderen Begriffen" – nämlich mit Stilbegriffen – als die historische Betrachtung ('historische' Begriffe). Die auch die kulturwissenschaftliche Betrachtung charakterisierenden Stilbegriffe, die "dauernde Tendenzen der Gestaltung" herausheben, machen es erforderlich, daß die Form-Analyse jeweils auf eine bestimmte "allgemeine Sinn-Funktion" ("geistige Energie") zurückgeht,[49] um die "Grundrichtung[en] symbolischer Formung, 'bildenden' Tuns überhaupt" zu erfassen. Bereits die Form-Geschichte erlaube es, sich partiell der "'reinen' Kontemplation" zu bedienen, die der "Welt der Formen" angemessen sei.[50]

Im Ms. 'Geschichte' finden sich auch schon einige Hinweise auf Parallelen bezüglich des Formbegriffs – und seines Kampfes mit dem Kausalbegriff, dem Zweckbegriff und dem substantiellen Formbegriff - in Geschichte und Biologie.[51] So sei "die Unsicherheit über das Verhältnis/Ursache – Form – Zweck/ [...] für die Erkenntnistheorie der Biologie wie für die der Geschichte das eigentliche [...] [Erkenntnis-]Hindernis." Beide seien gehalten, "die Eigentümlichkeit ihres

[46] Vgl. Cassirer, *Geschichte,* 176.
[47] Vgl. Cassirer, *Geschichte,* 177.
[48] Vgl. Cassirer, *Geschichte,* 232.
[49] Vgl. Cassirer, *Geschichte,* 177.
[50] Vgl. Cassirer, *Geschichte,* 217.
[51] Vgl. Cassirer, *Geschichte,* 68f.

Formbegriffs (= Sinnbegriffs) [zu] erfassen [und] kraft desselben den einseitigen 'Kausalismus' (den Begriff der 'mechanischen' Kausalität) [zu] überwinden." Die *Kategorie* Form sei in beiden Wissenschaften "als sui generis" zu erkennen bzw. anzuerkennen und "in ihrer Bedeutung (Funktion) der Kausalkategorie" gegenüberzustellen.[52] Ohne "der Eigenart und dem Eigenrecht der Kategorie der Form" gerecht zu werden, sei "der Aufbau der biologischen wie der geschichtlichen Welt nicht vollziehbar."[53] Aber obwohl die Geschichte seit Ranke die "ideelle Denkweise," die Goethe als erster in die Biologie eingeführt habe, anwende, und obwohl die Methode der Morphologie – im Anschluß an Goethe und Spengler – aus der Biologie in die geschichtliche Formbetrachtung übernommen worden sei, differiere aber das "Erkenntnisideal" der "Welt der Geschichte" nicht nur erheblich von dem "der Mathematik und [dem] der mathematischen Naturwissenschaften," sondern "auch von dem der biologischen Wissenschaften."[54]

b. Kulturwissenschaft als Formwissenschaft

Die Kultur des Menschen ist, so Cassirer im Ms. 'Geschichte,' an die "Wirksamkeit der 'Form' gebunden," weil sich in ihr kulturelles Wirken vollzieht und sich an ihr beständiges Umschaffen bewährt.[55] Der Mensch lebt sein Leben grundsätzlich "in dem allgemeinen Medium der Kulturformen," ein Heraustreten aus ihnen ist ihm verwehrt.[56] Diese Formen sind Gegenstand – und Erkenntnismittel – sowohl der Kulturwissenschaft als auch der Kulturphilosophie. Das "Eigentümliche d[ies]er Formbegriffe," die oft als Stilbegriffe auftreten, dürfe weder mit historischen Begriffen und Wertbegriffen (Windelband, Rickert) noch mit Gesetzesbegriffen konfundiert werden.[57] Die Kulturwissenschaft, so Cassirer 1942, zielt nicht auf "die Universalität der Gesetze" ab, aber ebensowenig auf "die Individualität der Tatsachen und Phänomene." Was sie erkennen will, ist "die Totalität der Formen, in denen sich menschliches Leben vollzieht."[58]

Die Vorlesung über 'Probleme der Kulturphilosophie' von 1939, die der Einsicht folgt, daß der Zugang zur Welt der Kultur sich allein über den ihr "eigentümlichen Formbegriff" eröffne,[59] hebt mit der Feststellung an, daß sich der

[52] Vgl. Cassirer, *Geschichte,* 126.
[53] Vgl. Cassirer, *Geschichte,* 242.
[54] Vgl. Cassirer, *Geschichte,* 147.
[55] Vgl. Cassirer, *Geschichte,* 204, 209.
[56] Vgl. Cassirer, *Geschichte,* 196f.
[57] Vgl. Cassirer, *Geschichte,* 231, 236.
[58] Vgl. Cassirer, "Zur Logik," 434.
[59] Vgl. Cassirer, *Kulturphilosophie*, 58.

"allgemeine Formbegriff" für jeden Gegenstandsbereich der Welterfahrung besondert, also auch für den der Kultur. Den "Problem- und Gegenstandstypus" der Kultur charakterisiere ein eigener konstitutiver "Auffassungsmodus," der, wie wir schon wissen, in der Ausdruckswahrnehmung im Unterschied zur Dingwahrnehmung seine Basis hat. Aber auch die "Kulturform" besondert sich in die Formen einzelner kultureller Bereiche wie Sprache oder Mythos. In jeder dieser Besonderungen enthält die Form ein qualitativ anderes Aufbauprinzip, Prinzip der Über- und Unterordnung.[60]

Allerdings sind Kulturobjekte auch in die Welt der Naturobjekte "eingebettet," "verkörpern" sie sich doch in Dokumenten, Monumenten, Denkmälern, weshalb sie gleichzeitig noch der kausalgesetzlichen Naturerkenntnis und der historischen Betrachtung unterliegen. Diese Erkenntnisweisen vermögen sie aber nicht zu erschöpfen bzw. ihren spezifischen Kultursinn zu erklären, was nur die kulturwissenschaftliche Formbetrachtung vermag. Für kulturellen Sinn sei charakteristisch, daß er symbolisch vermittelt ist und kein unmittelbares Dasein führt.[61] Kulturgegenstände "stellen etwas dar," "bedeuten etwas," "drücken etwas aus," dies aber in individueller Prägung durch die psychische Konstitution des sie schaffenden – und rezipierenden – Menschen. In die Definition eines Kultur-'Werkes' nimmt Cassirer diese drei Aspekte auf: das 'Werk' ist charakterisiert durch die Dimensionen des "physischen Daseins," das Darstellungsfunktion hat, des "Gegenständlich-Dargestellten," d.h. des dargestellten Sinns, und des "Persönlich-Ausgedrückten," der psychischen Eigenart des Darstellers.[62] Wenig später weist er noch zusätzlich auf das soziale Eigenleben des die Werke schaffenden Individuums als ebenfalls Kultursinn prägend hin.

Wenn das Spezifische, das die Kulturwirklichkeit ausmacht, im Kulturobjekt angelegter 'Ausdruck von Bedeutung' ist, dann ist die Ausdruckswahrnehmung bzw. das Ausdrucksverstehen diejenige Funktion, die ihn "ursprünglich originär" erfaßt.[63] Deshalb lautet eine der zentralen wissenschaftsphilosophischen Aussagen Cassirers: "Die Ausdruckswahrnehmung [ist der] natürliche [...] Ausgangspunkt aller kulturwissenschaftlichen Forschung."[64] Sie bilde ihr "Basisphänomen," während die "Sinneswahrnehmung" als Basisphänomen der Naturwissenschaft fungiert. Es sind die Form- und Stilbegriffe, die die Objektivation der

[60] Vgl. Cassirer, *Kulturphilosophie*, 59, 63, 66.

[61] Vgl. Cassirer, *Kulturphilosophie*, 69.

[62] Vgl. Cassirer, "Zur Logik," 400.

[63] Vgl. Cassirer, *Kulturphilosophie*, 69.

[64] Vgl. Ernst Cassirer, *Ziele und Wege der Wirklichkeitserkenntnis*, in ders. *Nachgelassene Manuskripte und Texte*. Bd. 2, Hrsg. von K. Ch. Köhnke und J. M. Krois (Hamburg: Meiner, 1999) [ECN 2], 168.

Ausdruckswahrnehmung ermöglichen, vollziehen. Im Ausdruckserlebnis erfahren wir das einzelne Kulturobjekt als "'Repräsentant,' Ausdruck eines Ganzen, das wir unmittelbar in ihm *vergegenwärtigen.*"[65] Die Form, Stil- und Gestaltbegriffe geben jeweils an, wie Einzelphänomene bestimmten Formen bzw. Gestalten zugeordnet werden. Die "gesamte Kulturwissenschaft besteht zuletzt in der Gewinnung solcher [...] Begriffe, durch deren [...] Anwendung wir ein individuelles Gebilde bestimmen," indem wir es einer Epoche, einer Kultur, einem Künstler zuordnen.[66]

Ein weiteres Spezifisches der Kulturwissenschaft wird an ihrer Lösung eines Grundproblems sichtbar, das sie mit der Naturwissenschaft teilt: am Problem, das Verhältnis von Besonderem und Allgemeinem zu bestimmen.[67] Cassirer erklärt die Beziehung von Individuellem und Allgemeinem sogar zum "Lebensfaden des Begriffs," der niemals zerschnitten werden dürfe.[68] Wenn für die naturwissenschaftliche Betrachtung "das Allgemeine in der Form eines 'Gesetzesbegriffs'" sich das Besondere subsumiert, so ordne die kulturwissenschaftliche "das Besondere dem Allgemeinen in [einer anderen – C.M.] Weise ein [...]."[69] Die bestimmt Cassirer als "Zusammenschau," die vom Einzelnen, Besonderen ausgeht, an dem sie das Ideelle (Form) schauend erfaßt, als ein "das Allgemeine *im* Besonderen, das Besondere *im* Allgemeinen anschauen."[70] Das Besondere repräsentiert, symbolisiert das Allgemeine (Goethe), das hier in "seiner konkreten Gestaltung" genommen wird, als ein "Inbegriff von 'Formen,'" nicht als ein System von Gattungsbegriffen.[71] Die "logische Arbeit sui generis" der kulturwissenschaftlichen Begriffe besteht im "Charakterisieren," nicht aber im "Determinieren."[72]

Die Kulturgegenstände sollten deskriptiv als ein Ganzes von Formen erfaßt werden, ohne daß metaphysisch nach dem Ursprung der Formen (Hegel, Spengler) gefragt werde könne, hinter das "Sein der Form überhaupt" könne grundsätzlich nicht zurückgegangen, nicht zurückgefragt werden.[73] Philosophen und Kulturwissenschaftler habe lediglich zu interessieren, 'was' die jeweiligen Formen bedeuten, ausdrücken. Durch "universale Überschau," so Cassirer, löse die konkrete Methode einer Kulturwissenschaft "aus der Fülle der Einzelphänomene ein 'Urbildliches' und Typisches" heraus, was "eine eigene und legitime Art der

[65] Vgl. Cassirer, *Kulturphilosophie*, 104.
[66] Vgl. Cassirer, *Kulturphilosophie*, 168f.
[67] Vgl. Cassirer, *Kulturphilosophie*, 133.
[68] Vgl. Cassirer, "Zur Logik," 427.
[69] Vgl. Cassirer, "Zur Logik," 428f.
[70] Vgl. Cassirer, *Kulturphilosophie*, 190.
[71] Vgl. Cassirer, *Ziele*, 174.
[72] Vgl. Cassirer, "Zur Logik," 431.
[73] Vgl. Cassirer, *Kulturphilosophie*, 97.

kulturwissenschaftlichen Begriffsbildung" darstelle.[74] Die kontemplativ-deskriptive "Phänomenologie der Formen" solle aber ebenso auch, wie schon die historische Betrachtung, die Formübergänge (Umbildungen, Metamorphosen), und damit ihre ideelle Entwicklung, wie ebenfalls die Korrelationen zwischen den Formen in den Blick nehmen.

Die Formen der Kultur zeichnen sich durch das bereits erwähnte Paradox von Permanenz und Fortwirkung aus, was ihre "Selbst-Entfaltung" bzw. Umbildung ermöglicht.[75] Obwohl der Lebensausdruck im objektiven kulturellen Formgebilde gewissermaßen "erstarrt," tragen diese "'erstarrten' Gebilde" weiterhin ein "eigentümliches 'Leben'" in sich. Sie besitzen das metamorphosische Vermögen, eine Füller neuer, einander ähnlicher Gestalten "aus sich hervorgehen zu lassen." Hierin, in der 'Lebendigkeit' der Form und in der Metamorphose der Formgestalten zeige sich "auch das Analogon, das immer wieder zu einem Vergleich der Kulturobjekte mit Objekten der organischen Natur geführt hat – / Diese Fähigkeit zur Permanenz der Form und zur Entwicklung der Form ist beiden gemein."[76] Allerdings legt Cassirer großen Wert auf die methodologische Einschränkung, wonach "die Kulturformen [...] 'Organismen' nur im *Als-Ob* Sinne" seien, und nicht im buchstäblichen Sinne. Das Übertragen biologischer Termini bzw. Charakteristika auf kulturelle und geistige Sachverhalte lehnt er kategorisch ab.

Das paradoxe Vermögen der Kulturformen nennt er auch ihre Prägnanz, weshalb er von "'geprägten Form[en]'" spricht, "die lebend sich entwickel[n]," die "sich ständig um- und weiterentwickeln."[77] Wandlungsfähigkeit der Form und Permanenz seien "die beiden entgegengesetzten und einander in dieser Entgegensetzung zugeordneten Pole des gesamten Kulturprozesses."[78] Vollzogen wird diese 'lebende Entwicklung' nicht durch irgend eine geheimnisvolle Kraft, sondern "durch die Arbeit der Individuen [...], die sich *in* diesen Formen ausdrücken."[79] "Jeder neue Gebrauch der Form" bedeute "zugleich eine neue 'Beseelung' der Form," bedeute "ein Einströmen neuen seelischen 'Lebens' in sie."[80] Das, was "individuelle Aneignung und individuelle Formung" leisten, nennt Cassirer eine Regeneration der Formen. Auch sie habe ihr Analogon in der organischen Natur, in der der individuelle Gebrauch von Funktionen die "morphologische Form" der Gattung verändere, während im geistig-kulturellen

[74] Vgl. Cassirer, *Ziele*, 162.
[75] Vgl. Cassirer, *Kulturphilosophie*, 126.
[76] Vgl. Cassirer, *Kulturphilosophie*, 127.
[77] Vgl. Cassirer, *Kulturphilosophie*, 128, 134.
[78] Vgl. Cassirer, *Kulturphilosophie*, 128.
[79] Vgl. Cassirer, *Kulturphilosophie*, 134f.
[80] Vgl. Cassirer, *Kulturphilosophie*, 131.

Leben "die Funktion stetig und unmerklich das Organ" verwandele, das so zu neuen Funktionen befähigt wird.[81] "Die allgemeinen, 'starren' Formen werden immer wieder in diesen Schmelztiegel des individuellen Gestaltens, Umformens zurückgeworfen" und entstehen aus ihm "in neuer vollkommener Gestalt."[82] Die Kulturwissenschaft habe sowohl die Art zu erforschen, wie "das individuelle 'Wirken' sich zu [bestimmten – C.M.] Form-Gebilden verdichtet," als auch "die Korrelation zwischen diesem 'Wirken' und diesen 'Gebilden'."[83] Indem sie aufweist, wie individuelles Wirken sich in einer universellen "Form ein-bildet," "in ihr und durch sie erfolgt," stoße sie auf die "Ur-Relation" allen geistigen und kulturellen Lebens. Weil "alles individuelle 'Wirken'" sich in eine "'vorgegebene' Form" einbildet, lasse sich in der Kulturwissenschaft das Kausalproblem nicht losgelöst vom Formproblem stellen. Die Form, die dem individuellen Wirken nicht begriffsrealistisch vorgegeben ist, kann sich selbst "nicht anders manifestieren [...] als in einer sich fortzeugenden Gesamtheit von [individuellen – C.M.] Taten."[84] Jegliches "Kulturphänomen [...] als individuelles Phänomen" besitzt seine Bedeutsamkeit allein als ein Schritt "im Werden 'zur' Form," und damit als Schritt der Objektivierung der Form.[85]

Die einzelnen kulturellen Formen lassen sich nun auf "gewisse 'Urgestalten', 'Grundformen'" zurückführen, auf die 'symbolischen Formen' der Kultur. Von denen habe die Philosophie zu zeigen, wie sie "generell zu definieren sind," die Kulturwissenschaft aber, "wie sie sich empirisch entfalten,"[86] denn das 'Formgesetz der Kultur', der in allen Grundformen auftretende "Rhythmus der Um- und Fortbildung," lasse "sich nicht abstrakt formulieren," sondern nur empirisch aufweisen. Nachdem die Kulturwissenschaft sich – mit Hilfe der Philosophie - des "objektiven Inventars der Kulturformen versichert" hat, geht sie mit Hilfe der Kulturgeschichtsschreibung daran, "uns das ganze reiche Gebiet" der jeweiligen Form "im weitesten Sinn" zu erschließen, indem sie "Vergangenes unmittelbar 'sichtbar'" macht.[87] Die Form- bzw. Gestaltbegriffe der Kultur "müssen [also – C.M.] mitten in das historische Werden gestellt, an diesem erkannt und beglaubigt werden." Deshalb kann Cassirer feststellen: "Kultur-Wissenschaft und Kultur-Geschichte" bedingen sich wechselseitig.[88]

---

[81] Vgl. Cassirer, *Kulturphilosophie*, 136f.
[82] Vgl. Cassirer, *Kulturphilosophie*, 139.
[83] Vgl. Cassirer, *Kulturphilosophie*, 187.
[84] Vgl. Cassirer, *Kulturphilosophie*, 190f.
[85] Vgl. Cassirer, *Kulturphilosophie*, 192.
[86] Vgl. Cassirer, *Kulturphilosophie*, 129.
[87] Vgl. Cassirer, *Kulturphilosophie*, 132.
[88] Vgl. Cassirer, *Kulturphilosophie*, 134, 143.

c. Biologie als Formwissenschaft

Nach Cassirers Verständnis hat sich im 20. Jahrhundert ebenfalls in der Biologie die "primäre Bedeutung des Faktors 'Form'" endgültig durchgesetzt.[89] Auch in dieser Wissenschaft stehen Formprobleme, die sich nicht in ein Verhältnis von Ursache und Wirkung auflösen lassen, als unableitbare an. Diese Tatsache lasse einige Analogien zur Kulturwissenschaft zu Tage treten. Mit Blick auf letztere formuliert er die Einsicht, wonach "schon in der Biologie [...] sich Form [bzw. Gestalt – C.M.] keineswegs restlos auf 'Gesetze' (ursächliche 'Kräfte') zurückführen" läßt, "schon hier ist sie ein Begriff sui generis." Das gelte auch umgekehrt, so daß beide Typen von Begriffen "ständig aufeinander bezogen werden müssen."[90] Die Gesetze für das organische Geschehen hätten die Form "heuristischer Maximen." Die Biologie werde allerdings in höherem Maße als die Kulturwissenschaft von kausalen Gesichtspunkten bestimmt, könne jedoch nicht ausschließlich durch sie aufgebaut werden.[91]

Nicht erst in den Kulturwissenschaften, sondern bereits in der modernen Biologie stehen die Formprobleme für "reine Strukturverhältnisse," für "Ordnung" und "System," und damit für eine spezifische Weise des Allgemeinen.[92] In der Biologie seien auch schon "morphologische Begriffe" unentbehrlich geworden,[93] wobei die Morphologie hier "als eine selbständige Grundwissenschaft [erscheint], die in der bloßen 'Entwicklungsgeschichte' nicht aufgeht / - ohne die vielmehr die Entwicklungsgeschichte selbst nicht durchführbar wäre."[94] Diese morphologische Methode sei nun auch auf die Kulturwissenschaften zu übertragen, denn "auch sie bedürfen durchaus des morphologischen Unterbaus." Morphologische Erkenntnis, also Form-Analyse, gilt Cassirer als ein begriffliches, theoretisches Erfassen der jeweiligen Form im Gegensatz zu ihrem physiognomischen, intuitiven Erleben (Spengler).[95] In der modernen Biologie habe in Bezug auf die "biologische Gestalt" bzw. das "Strukturgesetz" einer Lebensform zudem der – Goethesche – Gesichtspunkt des Formwandels, der Metamorphose endgültig Platz gefaßt.[96] Dabei faßt er

[89] Vgl. Cassirer, *Kulturphilosophie*, 94.

[90] Vgl. Cassirer, *Kulturphilosophie*, 165.

[91] Vgl. Cassirer, *Kulturphilosophie*, 189.

[92] Vgl. Ernst Cassirer, *Das Erkenntnisproblem in der Philosophie und Wissenschaft der neueren Zeit 4. Bd.: Von Hegels Tod bis zur Gegenwart (1832-1932). Gesammelte Werke.* Bd. 5, Hrsg. von B. Recki (Hamburg: Meiner, 2000) [ECW 5], 151, 250.

[93] Vgl. Cassirer, *Kulturphilosophie*, 157.

[94] Vgl. Cassirer, *Geschichte*, 232.

[95] Vgl. Cassirer, *Geschichte*, 233.

[96] Vgl. Cassirer, *Kulturphilosophie*, 127, 159. Vgl. Christian Möckel, "Formenschau, Formenwandel und Formenlehre. Zu Goethes Morphologie- und Metamorphosenlehre," *Goethe-Jahrbuch. (Goethe-Gesellschaft in Japan)* 52 (2010): 45-73.

Metamophose als "ideelle Genese," nicht als "historische Abstammung" auf.[97] "Die Biologie, so Cassirer 1942, hat den Zustand der bloßen Deskription und Klassifikation der Naturformen überschritten und ist zu einer echten Theorie der organischen Formen geworden."[98]

4. Resümee: Der Zusammenhang von Kausal- und Formbetrachtung

In Cassirers Wissenschaftsphilosophie stehen Kausalität und Formbetrachtung, wie schon mehrfach angesprochen, in keinem hierarchischen Verhältnis zueinander.[99] Beide wissenschaftlichen Betrachtungsweisen weisen einen eigenen Rechtsgrund bzw. einen unableitbaren urphänomenalen Charakter auf.[100] Sie bilden für eine isolierende Betrachtung zunächst bloße 'Gesichtspunkte,' ein bloßes regulatives 'Als-Ob,' sind dabei aber "beide nichtsdestoweniger 'objektiv,' konstitutiv, gegenstandsbestimmend." Diese "allgemeine Objektivierungsfunktion" realisiert sich auf verschiedene Weise: entweder werden Wahrnehmungs-Erlebnisse objektiviert ("aus-gesagt") oder Ausdrucks-Erlebnisse, entweder werden diese "unter Gesetze gefaßt" oder Form-, Stil- und Gestaltbegriffen zugeordnet.[101] Cassirer kann auch deshalb von einem "überraschend-einheitlichen Aufbau von Naturwissenschaft und Kulturwissenschaft" sprechen, weil beide auf das "Besondere und das Allgemeine als [...] korrelative Momente" abzielen. Auch dieses einheitliche Ziel wird unterschiedlich erreicht, in der Physik mit Hilfe des "Instrument[es] des 'Gesetzes', in der Biologie und Kulturwissenschaft [mit Hilfe - C.M.] der Form ('Gestalt')."[102] Die Objektivierungsfunktion erfordert es aber, daß jede der beiden Betrachtungsweisen "sofort durch das andere Moment ergänzt" wird, wobei diese Ergänzung, die sich in jedem Typus von Wissenschaft anderes darstellt, "rein umkehrbar [ist] - jedes Moment [...] das andere [verlangt,] um wirklich 'objektiv' (erfahrungs-bestimmend) zu werden."[103] Als fatal deutet er alle Versuche, in der Geschichtswissenschaft und in der Biologie, die Form als eine alternative Ursache einzuführen, so daß sich Form und Ursache auf ein und derselben Ebene begegnen. Dies führe in der Biologie in die methodischen Probleme des Vitalismus ("Dominanten") und in der Historie in die Auffassung der "Ideen" als "schaffender Kräfte."[104]

---

[97] Vgl. Cassirer, *Das Erkenntnisproblem*, 172.
[98] Vgl. Cassirer, "Zur Logik," 391.
[99] Vgl. Cassirer, *Geschichte*, 237f., 241.
[100] Vgl. Cassirer, *Kulturphilosophie*, 164.
[101] Vgl. Cassirer, *Kulturphilosophie*, 144.
[102] Vgl. Cassirer, *Kulturphilosophie*, 164f.
[103] Vgl. Cassirer, *Geschichte*, 239.
[104] Vgl. Cassirer, *Geschichte*,126f.

Die neue Sichtweise habe Eingang auch in die moderne Naturwissenschaft gefunden, die inzwischen zwei "autonome Fundamente sui generis" anerkennt, die "als wechselseitig sich haltende 'Potenzen'" in einem "Ineinander, Gleichgewicht" miteinander stehen.[105] So wie die "Ausschaltung des Formbegriffs" zugunsten des Kausalbegriffs aus der Wissenschaft und Philosophie seit dem 15. Jahrhundert Natur- und Kulturwissenschaften auseinanderriß,[106] führe seine Wiederentdeckung in nahezu allen Wissenschaften beide Pole des "Weltbegreifens" erneut zusammen. Nunmehr führe die Entwicklung in der modernen Physik, aber auch in der Biologie und in der Psychologie, Ganzheiten und Strukturen als "etwas Ursprüngliches, Unableitbares" für die Betrachtung ein.[107] Weil die Anerkennung des Form-, Ganzheits- und Strukturbegriffs "den Unterschied zwischen Naturwissenschaft und Kulturwissenschaft [aber – C.M.] keineswegs verwischt oder eliminiert," sondern nur "eine trennende Schranke" zwischen ihnen beseitigt hat,[108] müsse die Kulturwissenschaft dennoch "ihre Formen, ihre Strukturen und Gestalten" erforschen.

Die Gegenstandsbereiche der Wissenschaften – Physik, Biologie, Psychologie und Kultur (ideelle Sphäre) – scheinen für Cassirer eine Art Stufengang des Seins und seiner Erkenntnis zu bilden, wobei die Übergänge vom niederen Bereich zum höheren die niederen Gesetze bzw. Prinzipien in Geltung lassen und neue hinzufügen. Der Übergang vollzieht sich als qualitativer Sprung zu einem neuen "Problem- und Gegenstandstypus." Demnach würde der Gegenstandbereich der Kulturwissenschaften Gesetze bzw. Strukturen der Biologie ("Ganzheits-Bezogenheit") weiterhin einschließen, bereichert um ein spezifisches 'Kennzeichen' der Kultur. Für jeden Gegenstandsbereich bzw. jede 'Stufe' von Wissenschaft stehen die Gesetzes- und Formprobleme in einem anderen eigentümlichen Verhältnis.[109] In der Biologie bestehe "noch das Verhältnis, daß Gestaltbegriff und Gesetzesbegriff sich zwar nicht aufeinander zurückführen lassen – [...] wohl aber ständig aufeinander bezogen werden müssen."[110] In diesem Sinne ist für die Biologie ein "'Gleichgewicht' zwischen Formbegriffen und Gesetzesbegriffen" typisch. Die theoretische Physik wäre, wenn ich Cassirers Überlegungen fortspinne, durch den Primat der Gesetzesbegriffe gegenüber dem Formbegriff charakterisiert, während in den Kulturwissenschaften der explizit formulierte "Primat der Formbegriffe"[111] gegenüber dem Gesetzesbegriff hervortritt.

---

[105] Vgl. Cassirer, *Geschichte*, 240.
[106] Vgl. Cassirer, "Zur Logik," 449.
[107] Vgl. Cassirer, *Kulturphilosophie*, 57.
[108] Vgl. Cassirer, "Zur Logik," 455.
[109] Vgl. Cassirer, *Kulturphilosophie*, 92.
[110] Vgl. Cassirer, *Kulturphilosophie*, 165.
[111] Vgl. Cassirer, *Kulturphilosophie*, 166.

SKEPTICISM AND VARIETIES OF TRANSCENDENTAL ARGUMENT

Hamid VAHID

ABSTRACT: Transcendental arguments have been described as disclosing the necessary conditions of the possibility of phenomena as diverse as experience, self-knowledge and language. Although many theorists saw them as powerful means to combat varieties of skepticism, this optimism gradually waned as many such arguments turned out, on examination, to deliver much less than was originally thought. In this paper, I distinguish between two species of transcendental arguments claiming that they do not actually constitute distinct forms of reasoning by showing how they collapse into more familiar inferences. I then turn to the question of their epistemic potentials which I argue to be a function of both their types as well as their targets. Finally, these claims are reinforced by uncovering links between certain recent claims about the efficacy of transcendental arguments and the so-called Moore's paradox.

KEYWORDS: transcendental arguments, skepticism, situated thought, Moore's paradox

On the surface, at least, there does not seem to be much disagreement among philosophers about the salient features of what are known as Kantian transcendental arguments. These include their primary function (refuting varieties of skepticism), their subject-matter (experience and thought) and their structure (an extended modus ponens argument whose conditional premises are supposed to express necessary conditions for the obtaining of their subject matter). Kant's idea in proposing transcendental arguments was to identify the necessary conditions of the possibility of experience which he took to include the truth of the propositions that the skeptics generally seek to deny. Recent years have however witnessed major shifts of emphasis in regard to the nature of these features. Many theorists have cast doubt on the anti-skeptical potentials of transcendental arguments. Some have broadened their scope to include the possibility of language, self-knowledge and the like. And, as yet, there is no consensus about the status of the modality of their conditional steps.

In this paper, I begin by distinguishing between two species of transcendental arguments claiming that they do not actually constitute distinct forms of reasoning by showing how they collapse into more familiar inferences. This would then pave

the way for examining their epistemic potentials which will be argued to be a function of both their type as well as their targets. Finally, these claims will be reinforced by uncovering surprising links between certain recent claims about the efficacy of transcendental arguments and the so-called Moore's paradox.

1. Varieties of Transcendental Argument

As already noted, transcendental arguments are, formally speaking, extended modus ponens inferences with their conditional steps consisting of premises expressing necessary conditions for the obtaining of certain states such as experiences, thoughts, etc.

$$
\begin{array}{l}
p \\
p \rightarrow q \\
q \rightarrow r \\
. \\
. \\
* t
\end{array}
$$

Here p is supposed to express certain facts that even the skeptic ought to accept while the conclusion t is something that he typically denies. Kant, it seems, thought that some of his arguments in the Critique which are of such a type establish and reveal certain truths about the world, e.g., that we can have experience only if reality is indeed causal, that perception of change is possible only if there are persisting substances and so on. A number of theorists have, however, found transitions from psychological premises in these arguments to conclusions about an objective, non–psychological world puzzling and invalid. Barry Stroud, for example, has claimed that this subjective/objective gap can only be bridged with the help of some extra assumptions (like the verification principle) which would in turn render transcendental arguments redundant.[1]

Stroud seems to think that the subjective/objective divide is something like the is/ought gap in ethics though he presents no general argument as to why closing the former is impossible. Thus, barring an argument, the possibility of such transitions remains wide open.[2] He does however muster some support for his claim by examining a number of (more recent) transcendental arguments arguing that their success does not really hang on the truth of their conclusions but only on our believing them,

[1] Barry Stroud, "Transcendental Arguments," *The Journal of Philosophy* 65 (1968): 241-56.

[2] For acknowledging this possibility, see, for example, Ralph Walker, "Induction and Transcendental Argument," in *Transcendental Arguments*, ed. Robert Stern (Oxford: Oxford University Press, 1999), 13-31.

and, indeed, Kant himself seems to have also suggested transcendental arguments that are merely intended to establish how and what we must think about the world.[3]

Instead of getting bogged down in the debate over which type of such arguments deserves to be called 'transcendental,' it would be more prudent to admit that there are indeed two species of transcendental arguments sharing the same surface structure, and then try to evaluate the epistemic significance of each species. We may thus say that just as there are objective and subjective readings of the formal axioms of the probability calculus, there are objective and subjective interpretations of the formal structure of transcendental arguments. Accordingly, while objective transcendental arguments seek to reveal *truths* about the world, the subjective variety is intended to show why certain *beliefs* are indispensable for having thoughts, experiences and so on. I call these first-order and second-order transcendental arguments respectively (reflecting the type of content their conclusions express). With this distinction in mind, I shall now proceed to show that, their surface structure notwithstanding, transcendental arguments are actually species of inference to the best explanation.

As noted earlier, transcendental arguments purport to show that some conditions (C) are necessary for the obtaining of certain states or properties (E) (such as experiences of change, causality, etc.). However, as it turns out, it seems that all that such arguments do is to show that those conditions (together with certain background assumptions) are sufficient for the obtaining of E without being uniquely so. For example, in the Transcendental Deduction, Kant argues that for experience to be possible, it must belong to a unified consciousness which, in turn, requires that those experiences be experiences of physical objects. By itself, this does not rule out alternative ways of securing such experiences. Insisting on the impossibility of such alternatives seems to say more about the limitations of our imagination than those of the reality.[4] At most what the Kantian transcendental reasoning does is to narrow down alternative ways of securing such experiences by seeing, in Strawson's words, "how [our hypotheses] stand up to attack."[5] Under these circumstances, it would be more reasonable to speak of the necessary conditions of experiences as being the most appropriate means (relative to a set of *available*, rather than possible, alternatives) of securing such states.

---

[3] Walker, "Induction," 14.

[4] See, for example, Stephen Korner, *Fundamental Questions in Philosophy* (London: Penguin, 1971) and Richard Rorty, "Transcendental Arguments, Self-reference, and Pragmatism," in *Transcendental Arguments and Science*, eds. Peter Bieri, Rolf-Peter Horstmann, Lorenz Kruger (Dordrecht: Reidel, 1979), 77-103.

[5] Peter Strawson, *The Bounds of Sense* (London: Methuen, 1976), 98.

To give a concrete example, consider Strawson's version of Kant's argument in the Deduction (often referred to as the 'objectivity argument') which proceeds by showing why a particular alternative to the realist claim, viz., the sense-datum hypothesis, fails to satisfy what he takes to be the desiderata of experience.[6] This is done by trying to demonstrate which of these alternatives tell a better story about (or provide an explanation of) the obtaining of experience. This, in turn, involves examining which alternative hypothesis fits better with our (independently justified) background assumptions about the subject matter under discussion. Accordingly, Strawson proceeds by unveiling a number of (purportedly) justified claims about the nature of experience, e.g., that our experience possesses such features as unity, conceptualization, self-ascription and so on. He then considers two competing hypotheses, viz., the realist hypothesis emphasizing the existence of mind-independent objects and the sense-datum hypothesis construing these objects as bundles of sense data to see which alternative provides a better explanation of those particular features of experience. He concludes that the type of experience that is borne out by the sense-datum hypothesis fails to possess the kind of unity that our experience displays. Thus understood, we may see the function of transcendental reasoning as painting as coherent a picture or story of the obtaining of experience and its specific features: "We have before us the materials of a transcendental drama … In the Transcendental Deduction the story is told, the explanation is given."[7]

This way of conceiving of transcendental arguments, however, encourages viewing them as species of inference to the best explanation (IBE) where a hypothesis is adopted from among a number of competing alternatives in virtue of providing the best explanation of the available data.[8] The so-called explanatory virtues that are said to guide our choice of the best hypotheses often include such properties as simplicity, parsimony, non-ad-hocness, etc. There is, however, a great debate as to how these notions should be unpacked. It has been more customary to view them as reflecting purely methodological and pragmatic principles implying nothing about the world. But this raises the worry as to why possession of such (pragmatic) virtues should make a hypothesis more likely to be true.

A much more reasonable account of the nature of such virtues has been recently advanced by Sober who, rejecting the pragmatic approach, claims that appeals to such properties are actually surrogates for stating background assumptions about

[6] Strawson, *The Bounds of Sense*.

[7] Strawson, *The Bounds of Sense*, 86.

[8] For an early suggestion along these lines see Jay F. Rosenberg, "Transcendental Arguments Revisited," *The Journal of Philosophy* 72, 18 (1975): 611-624, which views transcendental arguments as "no different from a Peircean abduction" (623) but does not substantiate it.

the subject matter and the inference problem one faces.[9] For it is only relative to a set of background assumptions, he says, that observations can be said to support one hypothesis better than another (there is obviously nothing sacrosanct about these assumptions as they themselves may be challenged). So the function that properties such as simplicity are intended to serve is to help bring observations to bear on hypotheses. Viewed in this light, the credibility of an IBE turns out to depend on how well the selected hypothesis coheres with the rest of our (independently justified) background assumptions and this accords very well with our diagnosis of transcendental arguments. We may, thus, conclude that a Kant-style transcendental argument is actually an IBE in disguise.[10]

By way of supporting this conclusion, consider a recent statement of some of Kantian transcendental arguments (due to Sacks[11]). His first example is taken from the First Analogy where Kant seeks to show that all change is merely the alteration of an abiding substance. On Sacks' reading, since our perception of change has a unified character, that is possible only if something persists across our experiences. This seems to be just another way of saying that the best way to account for (explain) the unified character of our relevant experiences is to postulate the existence of persisting objects. The second example is drawn form the Second Analogy where Kant tries to establish why everything that happens has a cause. According to Sacks, Kant's thought is that where one's perceptions are of an event, then the order of these perceptions is irreversible. The agent would then be faced with the puzzle of the irreversibility of two perceptions that can only be accounted for if the perceived states of affairs themselves happened one after another: "[I]t's only thus that the subject can … *explain* the irreversibility of his perceptions of them."[12] This is, again, another way of saying that the perceived irreversibility is best explained by postulating a corresponding relation within the objective domain.[13]

---

[9] Eliot Sober, *Reconstructing the Past* (Cambridge, Mass.: MIT Press, 1988).

[10] For a fuller defense of this suggestion see Hamid Vahid, "The Nature and Significance of Transcendental Arguments," *Kant-Studien*, 93, 3 (2002): 273–90. Here is a recent endorsement of this claim: "Kant's most important method, the transcendental method, is also at the heart of contemporary cognitive science. To study mind, infer the conditions necessary for experience. Arguments having this structure are called transcendental arguments. Translated into contemporary terms, the core of this method is inference to the best explanation, the method of postulating unobservable mental mechanisms in order to explain observed behavior." (Andrew Brook, "Kant's View of the Mind and Consciousness of the Self," *The Stanford Encyclopedia of Philosophy*, ed. Edward N Zalta (Summer 2007 Edition), http://plato.stanford.edu/entries/kant-mind/.)

[11] Mark Sacks, 2005, "The Nature of Transcendental Arguments," *International Journal of Philosophical Studies* 13, 4 (2005): 439-460.

[12] Sacks, "The Nature," 450.

[13] I have already emphasized that we must be content here with the best *available* explanations.

Before leaving this topic, it would be instructive to examine Sacks' own account of the nature of transcendental arguments which promises to explain a number of peculiar features of such arguments. Although I am going to challenge his account of transcendental arguments, nothing in what he says contravenes the conclusions we have reached so far. Indeed, as we shall see later, his view of the matter provides an interesting gloss on my own account of the epistemic efficacy of transcendental arguments. Sacks' main concern is with the status of the conditional or ampliative steps ("q→r") in transcendental reasonings. However, instead of taking their bare propositional content, "q→r," as constituting such steps, we should consider, he says, the "situated thought" of the content in question, viz., the thought that one would have from a particular point of view within a framework. For example, while the propositional content of q, say, "there is a tree in the garden," doesn't entail r, "there is a tree in front of me," the situated thought of q does. Consider an agent situated in front of a tree in a garden, thus, coming into perceptual contact with it. Under these circumstances, the agent can have the situated thought that there is a tree in the garden (s"q"). Being in possession of s"q," he can come to know, without further observation, that there is a tree in front of him.

Sacks' general idea is that while, to substantiate the conditional premises of a transcendental argument, an argument is needed to sanction the move from "q" to "r," the situated thought of "q" would suffice to license a priori the ampliative move in question. For instance, while it is difficult to see why the central claim in the First Analogy, "Every change is merely the alteration of an abiding substance," should be true, one can easily see how the move from the situated thought of something changing to the existence of a pertinent persisting object can make sense. The same goes for Kant's other claims like every event has a cause, and its accompanying transcendental argument. Here too, says sacks, the relevant situated thought would enable us to see why the claim in question is true.

Sacks' gloss on transcendental arguments raises a number of interesting questions. To begin with, the notion of the situated thought of a content does bear some resemblance to the phenomenalist analysis of the notion of an object (without its reductionist associations). Worried about the skeptical consequences of the realist claim that there is a genuine distinction between the conception we form on the basis of experience and the way things are, phenomenalists sought to close the appearance/reality gap by reducing statements about physical objects to those about our perceptions of them. Accordingly, a statement like q, "There is a tree in the garden," is to be reduced to those describing our sensory impressions when we come into perceptual contact with the tree in question. This seems to be akin to what the thesis of situated thought involves. A situated thought of q requires that

"we …. envisage a subject situated in front of a tree in the garden, and being perceptually related to it."[14] Moreover, just as phenomenalists qualified their initial thesis by admitting possible as well as actual perceptions in order to account for statements about unperceived physical objects, thus, analyzing, say, q in terms of such statements as "if I were to see such a tree before me, then certain sensory impressions would follow," Sacks also makes a corresponding distinction between actual and possible situated thoughts: "a situated *thought* differs from the corresponding experience in that the situated thought does not require that the subject actually be situated – only that he approximate in thought to what would be delivered up to him if he were so situated."[15]

An initial problem with the situated-thought thesis is that even if we can establish that there is a valid transition from the situated thought of the content expressed by q (s"q") to "r," why should it follow that there is likewise a valid transition from "q" to "r" when it is, in fact, "q→r" that appears as a premise in the pertinent transcendental argument? Another problem concerns the nature of a situated thought itself. Is it sufficient for a thought to be situated that an agent be perceptually related to the object of that thought? Sacks himself denies this by providing the following illustration. I am in a laboratory vehicle, receiving pictures on a screen from a camera that I know is in the garden, though I do not know that the vehicle is in the garden. This is not yet a case of a situated thought, for although my knowledge of q is based on my own perception, I fail to infer r from q. What is needed, he says, is that the thought in question be grasped from one's point of view and be informed by it. Only then "the mere having of the thought licenses the move to 'there is a tree in front of me.'"[16]

But it seems that these conditions still fail to be sufficient. To see this, consider the following example.[17] Suppose there is a candle in front of an agent but he doesn't see it directly. Rather, there is a series of elaborately disguised mirrors that reflect an exact image of the candle. These mirrors are placed next to one another in such a way that what the agent sees is the image of the candle in the last mirror in the series in front of him where the candle itself is and hiding it from his view. In this case, the agent's sensitive[18] and justified belief that he is looking at a candle is based on his own perception from which he can justifiably infer that

[14] Sacks, "The Nature," 445.

[15] Sacks, "The Nature," 444.

[16] Sacks, "The Nature," 445.

[17] Adapted from Gilbert Harman, *Thought* (Princeton: Princeton University Press, 1973).

[18] This belief is sensitive in the sense that if there were no candles in the room the agent would not believe that he is looking at one.

"There is a candle in front of me," thus, satisfying Sacks' strictures on a situated thought, but, intuitively, we do not want to say that the agent's thought is really situated.[19]

Let us now turn to the application of the situated-thought thesis to transcendental arguments. Sacks' idea, we may recall, is that only by taking conditional sentences in such arguments as expressing situated thoughts can there be any hope of establishing their conclusions. He gives, as an example, Kant's claim in the First Analogy, namely, that "All change is merely is the alteration of an abiding substance." It is difficult to see why this claim should be true, he says, but things will be different once we consider the situated thought involving its content.[20] But I do not see why we need to resort to the situated thoughts involving the changes, rather than the changes themselves, to substantiate the claim in question. Sacks himself provides a corresponding argument focusing not on the experiences but on the states of affairs apprehended: "For [the state of affairs] to be presented as changing in relation to one another, or as abiding, there must also be an abiding external world on the backdrop of which it makes sense to say either that where one object was there is now another, or that where one object was there is still the same object."[21] It may be that Sacks thinks ascending to a situated-thought level would provide an immediate a priori route to the claim that there are abiding substances just as the mere having of the situated thought that there is a tree in the garden immediately licenses the move to "there is a tree in front of me." But this does not seem to hold in the transcendental case discussed as Sacks provides an elaborate argument (filling a full paragraph) to show why the claim in question holds. Similar points apply to Sacks' other example concerning causation.

Finally, Sacks claims that what is both necessary and sufficient for a transcendental proof to work is only that the subject's thought be situated virtually. This claim seems to limit the scope of transcendental arguments only to perceivable subject matters. But this is unwarranted. There are no reasons why one cannot set up a transcendental argument to establish claims about unobservable entities or properties. A clear example of this is Sacks' own discussion of the principle of causation according to which every event has a cause. In a nutshell the claim is that the irreversible order of our perceptions of two events can only be explained if the perceived states of affairs are put together in a such a way that one state could not

[19] Maybe there is more to the notion of a situated thought that would enable it to accommodate such scenarios. But these need to be made more explicit.

[20] If we consider a change as constituting an event and construe events in the standard fashion as the exemplification of a property by an object, does this claim not turn out to be analytically true?

[21] Sacks, "The Nature," 448-9.

happened except after the occurrence of the other. But this argument at most establishes an objective, irreversible, order between events themselves. As such, it is quite consistent with a Humean regularity view of causation. That the objective relation is also necessary is neither here nor there. Nomologically necessary relations are widely thought to be unperceivable. Having identified varieties of transcendental arguments and delineated their structure, I shall now turn to the question of their epistemic effectiveness.

2. The Epistemological Significance of Transcendental Arguments

The epistemic potential of transcendental arguments is a function of both their type as well as their targets. So I begin with first-order transcendental arguments whose conclusions have first-order contents. Such arguments purport to show that the truth of certain statements is the necessary condition of the obtaining of certain types of states such as thoughts, experiences and so on. Let us start with the most well-known of such first-order transcendental arguments, namely, those used against Cartesian skepticism. Versions of this argument include Kant's argument in the Deduction and Strawson's variation on the same theme (the objectivity argument). I shall not directly challenge these arguments. Rather, I shall examine their legitimacy in the light of our construal of first-order transcendental arguments as species of IBE. Accordingly, what these arguments seek to show is that the realist hypothesis is a better explanation of our experiences and their specific features than other competing alternatives such as the Cartesian demon hypothesis, etc., in virtue of being simpler, more parsimonious and the like.

As regards the structure of IBE, it was noted that the explanatory virtues, in terms of which we are supposed to discriminate between competing hypotheses, should be seen as surrogates for stating background assumptions about the subject matter under discussion. Accordingly, the success of IBE depends on the contexts where such an argument is being deployed for some of these contexts possess such a low degree of epistemic involvement that they do not allow invoking the required assumptions, thus, hindering the process of adjudicating between the competing hypotheses. The context of establishing the realist (external world) hypothesis is precisely such a context. Here the skeptic claims that our having the experiences we take to be of physical objects (with all their specific features) is compatible with the possibility of there being no such objects at all while a demon (or a super-scientist) is inducing these experiences in us. No common grounds other than these experiences exist between the realist and the Cartesian skeptic. At a level this basic, we are, on pain of begging the question, precluded from appealing to background assumptions about the subject matter in question (be it the external world, the demon or whatever).

So if appeals to simplicity and other explanatory virtues are to be seen as involving background assumptions about a particular subject matter, IBE cannot be put to use to ground our preference for realism over, say, the Cartesian skeptical hypothesis. Now, if we take the view that first-order transcendental arguments are actually species of IBE, it follows that the corresponding anti-skeptical transcendental arguments cannot also be expected to deliver the goods. On the other hand, if the context of applying an IBE is one which possesses a high degree of epistemic involvement (as in scientific or ordinary contexts), then the invoking of the relevant background assumptions would not be unjustified because such contexts do not generally impose strict epistemic constraints on what can be said to be known. Accordingly, such inferences stand or fall depending on whether their pertinent background assumptions can be adequately defended.[22] The same holds for those transcendental arguments that collapse into such inferences.

Let us now turn to second-order transcendental arguments whose conclusions are some belief statements. Stroud, we may recall, claims that it is not the truth of certain propositions that is needed to explain the obtaining of experience, knowledge, and the like, but the belief in those propositions. As such, typical transcendental arguments cannot undermine radical skeptical positions. Recently, however, Stroud has claimed that these 'modest' transcendental arguments still possess anti-skeptical bite. A number of philosophers, who have taken the Stroudian line on the nature of transcendental arguments, have also tried to show that modest transcendental arguments possess significant anti-skeptical consequences. In what follows, I consider and reject two recent attempts in that direction before turning to Stroud's own argument.

Using Kant's (anti-Humean) response to the problem of induction as a case study, Ralph Walker has claimed that Kant's argument is not really intended to establish the reliability of induction but only that belief in such a hypothesis is indispensable if we are to have any thoughts at all.[23] He distinguishes between 'third-personal' and 'second-personal' transcendental arguments. The former aims at establishing certain truths about the world while the latter only aims to show that certain beliefs are indispensable for thought. So although Kant's transcendental argument for induction is of a second-personal variety, it is, he thinks, still effective against the skeptic as it justifies our inductive practices. I shall not dwell on Walker's contention about the typology of Kant's argument for induction, but only

[22] For a defense of this position see Hamid Vahid, "Realism and the Epistemological Significance of Inference to the Best Explanation," *Dialogue* XL (2001): 487-507.

[23] Walker, "Induction".

concern myself with assessing his claim that it justifies our belief in induction. I think the claim is unwarranted.

Although Walker admits that "there may also be transcendental arguments which are third-personally valid,"[24] he thinks they are generally problematic. In any case, Kant's argument for induction is not third-personal valid because he himself admits that he cannot show space and time to be the only possible forms of sensible intuition. But, taken second-personally, the argument, says Walker, is plausible because *our* experience has spatio-temporal character and so space and time must be a priori forms of intuition. As for non-spatio-temporal forms of sensible intuitions, Walker dismisses the possibility as irrelevant on the ground that "[p]eople who claimed not to have spatio-temporal experience would not be worth arguing with, and if their experience really *was* different we should be unable to recognize them as having experience, or concepts, or language at all."[25] Walker's reasoning is, however, problematic as his latter remarks actually constitute a third-personal transcendental argument (similar indeed to Davidson's well-known argument[26] against the possibility of radically different conceptual schemes). For what it says is that for people to be interpretable (i.e., being recognized as having experience, concepts, etc.), their experiences must be similar to ours. And since people are interpretable, the proposition "people's experiences have spatio-temporal structure" must be true. Thus, Walker's illustration of the problematic nature of third-personal transcendental arguments presupposes their validity.

Let us, however, return to his claim about the anti-skeptical potentials of Kant's argument for the principle of induction. An early sign of confusion appears when Walker introduces a distinction in order to elucidate what the argument is intended to achieve: "[R]efuting Hume's skepticism about induction would not be a matter of showing that our expectations will continue to be fulfilled. It would be a matter of showing that it is *rational to expect* that they will, not just as a matter of custom or habit. What Kant tries to do is to show by means of a transcendental argument that this is indeed rational."[27] However, if we take 'rational' in an epistemic, truth conducive sense, then Walker's distinction is no longer tenable, for to show that belief in induction is justified (rational) is to show that it is more likely to be true. Walker further elaborates the sense in which transcendental arguments are, in his view, rational or justification-conferring. The skeptic calls into question

[24] Walker, "Induction," 23.

[25] Walker, "Induction," 24.

[26] Donald Davidson, "On the Very Idea of a Conceptual Scheme," in his *Inquiries into Truth and Interpretation* (Oxford: Oxford University Press, 1984), 183-98.

[27] Walker, "Induction," 16.

whether we are entitled or justified to rely on induction. Walker thinks that "transcendental arguments, taken in the second-personal way, answer this. Certainly we are entitled to do so, and for the best possible reason: there is no serious alternative available to us."[28] He then considers the worry that second-personal transcendental arguments make no reference to truth, but goes on to say that third-personal transcendental arguments do no better.

But Walker's 'having-no-other-option' response hardly addresses the worry just raised. For what actually lies behind the objection is whether the type of justification purportedly provided by second-personal transcendental arguments has epistemic significance. To see what is at issue, we need to remind ourselves of various senses of justification (rationality). Justification is a goal-directed notion with different goals entering into different conceptions of justification. If one's goal is maximizing truth and minimizing falsity or achieving a comprehensive and accurate set of beliefs, then a belief serving this goal would be justified in an epistemic sense. But when our goal is, say, maximizing satisfaction of desire, then the ensuing concept of justification would be prudential. This gives rise to the possibility that one and the same belief might be justified for an agent in one sense but unjustified in another sense. For example, suppose it is rational to a high degree for a mediocre academic to believe that his believing that he is more talented than his colleagues will help him succeed at his job (by boosting his confidence). Then, all things being equal, it is practically rational for him to believe that he is more talented than his colleagues. But this very belief may be unjustified if viewed from the epistemic point of view of achieving an accurate belief system.

In general, we may say that if an agent has a certain goal G and if, on careful reflection, he would believe that M is an effective (or indispensable) means to G, then, all else being equal, it is rational for him to bring about M.[29] We can now see that Walker's argument regarding the indispensability of belief in induction, as a means to the obtaining of thought and experience, at best confers justification on the belief in question only in a non-epistemic sense because the goal is not the maximization of truth but the obtaining of certain psychological states. Thus viewed, it is hardly encouraging news for the realist to be told that "[s]econd-personal transcendental arguments are [his] only real defense against skeptical doubt."[30]

---

[28] Walker, "Induction," 23.

[29] See, for example, Richard Foley, *The Theory of Epistemic Rationality* (Cambridge Mass.: Harvard University Press, 1987). Note that Walker's claim cannot be that having a *justified* belief (in a truth conducive sense) in induction, as opposed to merely believing it, is necessary for the possibility of thought, otherwise his second-personal argument would collapse into a third-personal transcendental argument.

[30] Walker, "Induction," 28.

A rather different tack has been tried by Robert Stern who also thinks that second-order transcendental arguments (which he calls 'belief-directed' arguments) have anti-skeptical consequences.[31] He distinguishes between epistemic skepticism which requires certain knowledge in the face of skeptical doubts and justificatory skepticism which does not ask for conclusive reasons or justification but settles for much less, namely, reasons that are, by an agent's lights, justification-conferring. Stern thinks that this 'internal' conception of justification is what Sosa has called 'reflective' justification. Sosa identifies this species of justification with the outcome of the application of our deepest intellectual norms and a matter of perspectival coherence.[32] On this view, according to Stern, even the beliefs of the inhabitants of a demon world can be justified. He claims that Kant takes coherence to be one such norm so that if an agent's belief system is more coherent with the belief p than without it, then this belief is justified for that agent.[33] Second-order (belief-directed) transcendental arguments, says Stern, are justification-conferring precisely because they ensure that the adding of their conclusions to our belief system enhances its coherence.

Stern's account raises a number of issues. One can see his appeal to the coherence norm as a way of supplementing Walker's proposal. The main problem with Walker's approach, as we saw, was that his underlying conception of justification didn't really hook up with truth resulting, at best, in a non-epistemically justified belief. One can see Stern's appeal to his version of the coherence theory of justification as a way of restoring the link between justification and truth. He admits, however, that he cannot establish that his coherence norm is truth conducive but also denies that such "norms are thereby rubbed of their justificatory capacity, insofar as they seem to offer the best guide to truth from where we are."[34] But this remark hardly addresses the question of why his coherence-based conception of justification is truth conducive. And there are also the standard objections to the claim that mere internal coherence yields a reason for thinking that the beliefs of a system are true. These include the possibility of alternative, equally coherent systems; gaining epistemic access to the coherence of one's belief system (especially if coherence is not merely understood in a negative

[31] Robert Stern, "On Kant's Response to Hume: The Second Analogy as Transcendental Argument," in *Transcendental Arguments*, ed. Robert Stern, 47-67.

[32] See various articles in Ernest Sosa, *Knowledge in Perspective* (Cambridge: Cambridge University Press, 1991).

[33] This requires that we be able to determine whether the coherence of our belief system has been enhanced, an ability whose psychological feasibility is doubtful.

[34] Stern, "Kant's Response," 59.

sense but also as meaning that the beliefs in question must positively support one another).

Moreover, it is not clear that the truth-conducivity of this particular brand of the coherence norm is compatible with what Stern claims about the justified status of the beliefs held by the inhabitants of a demon world for such beliefs are, by hypothesis, systematically false. Indeed, Sosa's views on this issue are more complex than are made to appear here. Sosa makes a distinction between 'aptness' and 'justification.' An apt belief is one that is produced by a reliable or virtuous faculty in the environment in which it is operating. A justified belief, on the other hand, is construed in terms of the notion of an 'epistemic perspective' which is, in turn, cashed out as consisting of meta-beliefs concerning the faculty (responsible for producing the target belief) and its reliability. Commensurate with this distinction, Sosa makes another distinction between animal knowledge (which requires only apt belief) and reflective knowledge (requiring both apt and justified belief).

Sosa is not ideally clear about the relationship between the notions of 'justification' and 'aptness.' But it seems that he takes these notions to be independent of one another. An apt belief is not automatically justified, and a justified belief need not be apt. This interpretation is further supported by Sosa's claim that while animal knowledge involves only apt belief, reflective knowledge requires both apt and justified belief. The fact that both justification and aptness are required for reflective knowledge shows that neither implies the other. And this is as it should be for Sosa regards "justification [as] amount[ing] to a sort of inner coherence."[35] Sosa's response to the problem of the status of beliefs in demon scenarios thus involves appealing to the distinction between aptness and justification, but, in so doing, he relativizes assessment of the epistemic status of the victim's beliefs in the demon world: "Relative to the demon's [environment] D, the victim's belief may be inapt and even unjustified … Even so, relative to our environment…the beliefs of the demon's victim may still be both apt and valuably justified."[36] Justification-relativity does not seem to be congenial to Stern's concerns.

So given the preceding problems for Stern's coherence theory of justification, it is not clear whether he would wish to follow Sosa's lead in defending the truth-conductivity of his conception of justification. To avoid problems of exegesis, however, I shall propose a dilemma for Stern's account and argue that both horns of the dilemma undermine its viability. Let us begin by assuming that coherence is not truth conducive. In that case, Stern's account can hardly be seen as an improvement on Walker's indispensability argument. Both yield, at most, justified beliefs in a

[35] Sosa, *Knowledge in Perspective*, 289.

[36] Sosa, *Knowledge in Perspective,* 290.

non-epistemic sense that could hardly counter the threat of skepticism. Suppose however that Stern's brand of coherence theory is truth conducive – assuming that the standard objections to such a theory, especially the problem of epistemic access, have been satisfactorily dealt with. Then, transcendental arguments are rendered redundant. For to see whether a belief p is justified, all we need to do is to find out, as Stern puts it, whether our "belief-set is more coherent with the belief that p as a member than without it."[37] There will thus be no need to involve transcendental arguments. This creates a dilemma for Stern's account of transcendental arguments analogous to that posed by Stroud for the standard readings of such arguments. Either second-order transcendental arguments yield non-truth-conducive coherence in which case they lack anti skeptical potentials, or they generate truth conducive coherence in which case they will become redundant.[38]

Let us now turn to Stroud's own gloss on the epistemic efficacy of the second-order transcendental arguments. Although he rejects 'ambitious' first-order transcendental arguments, he thinks that the second-order variety still possesses some epistemic value in the face of the skeptical challenge, and illustrates this through the following second-order transcendental argument.[39] Suppose it has been established (as claimed by Strawson) that to think of a world independent of us we must think of it as containing enduring particular objects. What does this entail? Well, if the claim is true, it follows that we could not think of there being anything in the world (such as people) without thinking that there are enduring objects. This, in turn, entails that we could not think of these people as believing things and in particular as believing that there are enduring particulars without believing that there are enduring particulars. So we could not ascribe that belief to people without thinking that it is true. In other words, we cannot consistently find that people are wrong in believing that there are enduring particulars. This does not show that the belief in enduring objects is true but only that we cannot consistently deny it. As such the belief in question would be indispensable to our having any thoughts or beliefs about the world at all. Stroud takes this to imply that certain beliefs are invulnerable in the sense that "we could never see ourselves as holding the beliefs in question and being mistaken."[40]

Let us assume that Stroud is able to extend this treatment to all second-order transcendental arguments and draw the same conclusion, viz., that we could never

---

[37] Stern, "Kant's Response," 54.

[38] For further criticisms see Mark Sacks, "Transcendental Arguments and the Inference to Reality: A Reply to Stern," in *Transcendental Arguments*, 67-83.

[39] Barry Stroud, "The Goal of Transcendental Arguments," in *Transcendental Arguments*, 155-173.

[40] Stroud, "The Goal of Transcendental Arguments," 168.

see ourselves as holding the relevant beliefs and being mistaken. What does follow from this? The immediate upshot is to ensure the consistency of our belief system. Applying the relevant transcendental arguments, we are, in effect, implying that we can attribute certain beliefs to people (including ourselves) only if we already hold them to be true. This, at most, guarantees the consistency, not truth, of our beliefs. Indeed Stroud's argument is just an application of Davidson's general requirement on a correct theory of interpretation, viz., the principle of charity. According to Davidson to attribute thought and language to speakers we must assume that they hold similar beliefs as ours. Indeed, Stroud, himself goes on to illustrate his point by referring to the "closely related position of Donald Davidson who, from conditions of belief ascription, or what he calls 'radical interpretation,' concludes that most of our beliefs must be true."[41] Stroud, however, seems to think that Davidson holds the strong view that the application of charity guarantees the truth of most of our beliefs are true, and instead suggests the weaker conclusion that we cannot attribute beliefs to speakers which we see as mistaken. But this is consistent with Davidson's early formulations of charity which only require "assigning truth conditions to alien sentences that make native speakers right when plausibly possible, according, of course, to our own view of what is right."[42] For all we know, however, the interpreter's beliefs might very well be mistaken. So Davidson's principle of charity at most guarantees agreement (consistency) between the interpreter and the speaker. This means that the epistemic efficacy of second-order transcendental arguments goes only as far as establishing the coherence of a belief-set.[43]

The problem can be made more explicit once we consider Stroud's claim that beliefs borne out by second-order transcendental arguments are invulnerable. For 'invulnerability' is a non-truth-conducive feature of such beliefs and, as such, of little consequence for defeating skepticism.[44] When Stroud says, of such beliefs, that we could not see ourselves as holding them and being mistaken, he is only highlighting the fact that we would fall into a pragmatic (not epistemic) trap if we deny them. This is very much like the position we get ourselves into when we assert the so-called Moorean's sentences of the type "p is false but I believe that p," "~p & IBp," for to assert such a sentence is to assert ~p (i.e., to express the belief that p is mistaken) and yet go on to say that one believes that p. What is peculiar about

[41] Stroud, "The Goal of Transcendental Arguments," 166.

[42] Donald Davidson, "Radical Interpretation," *Inquiries into Truth and Interpretation*, 137.

[43] For further elaboration of this point see Hamid Vahid, "Charity, Supervenience and Skepticism," *Metaphilosophy* 32, 3 (2001): 308-325.

[44] It is for this reason that Davidson himself adds an externalist slant to the principle of charity to give it some epistemic bite. See Vahid, "Charity, Supervenience and Skepticism" for elaboration.

Moorean sentences is that while they have coherent truth conditions, their assertion lands us in pragmatic, not logical, contradiction. Our inconsistency arises not from what we are claiming but from the fact that *we* are claiming it.

This squares very nicely with the proposed role of the principle of charity in constituting the epistemic significance of second-order transcendental arguments. For once the principle of charity is recognized as being constitutive of intentional ascription, of what it is to be a speaker at all, we can see how the assertion of Moorean sentences undermines charity. Consider an instance of "p & ~IBp," say, "It is raining but I do not believe that it is raining." By asserting p, or assenting to "p," one is performing a speech act to communicate certain information that p. Thus, when an agent assents to "It is raining" the default position is to interpret the utterance in such a way that it is true just in case it is raining, i.e., take it to mean "it is raining." And, assuming Davidson's strictures on interpretation, to infer, in accordance with the principle of charity, that she believes that it is raining. So when, having assented to "it is raining," the speaker goes on to assert that she does not believe that it is raining, this would be a clear case in which the principle of charity is undermined. So it is because the assertion of Moorean sentences contravenes the principle of charity that these sentences are defective.[45]

Thus, when Stroud says that second-order transcendental arguments have epistemic value because one could not see himself as holding the belief in their conclusion p, say, "There are enduring objects," and being mistaken, he is merely saying that the sentence "~p & IBp" is Moorean in that its assertion is paradoxical. But the paradox is merely pragmatic and this, in turn, reflects the epistemic import of second-order transcendental arguments. As noted earlier, such arguments at best show that their conclusions are pragmatically, not epistemically, justified. As such they cannot be of much comfort to the anti-skeptic. To conclude, on any of the accounts discussed so far, transcendental arguments, of first-order or second-order variety, seem to lack significant epistemic bite.

[45] See Hamid Vahid, "Radical Interpretation and Moore's Paradox," *Theoria* 74 (2008): 146-163.

A PEACE PLAN FOR THE SCIENCE WARS

Mark Owen WEBB

ABSTRACT: In what has become known as the 'Science Wars,' two sides have emerged. Some philosophers of science have claimed that, because science is a social practice, it is hopelessly infected with political bias. Others have claimed that science is a special kind of practice, structurally immune to bias. They are both right, because they are referring to different things when they use the word 'science.' The second group is referring the method of theory selection, as practiced by scientists in the laboratory, while the first group is referring to the ongoing social practice of science, of which theory choice is a part. The scientific method of theory choice, when practiced correctly, is resistant to bias, while the socially embedded practice is particularly sensitive to political forces, and so is subject to bias.

KEYWORDS: science, scientific method, science studies, theory, explanation

I.

Most wars are about territory, and a peace plan usually consists of some way to give both sides as much of the desired territory as is possible consistent with other claims of justice. The two (or more) sides never all get all of what they want. The science wars are a bizarre variation on this theme, and my proposed peace plan will have to be correspondingly bizarre. First of all, the territory at issue here is territory in logical space; each party to the science wars seeks to stake out a position, defend a thesis, that appears to be inconsistent with the other side's position and thesis, driving the enemy into retreat. But things are not always as they seem. My claim is that the two sides are occupying different territory, which they mistakenly believe to be the same territory. It is as if, in response to Hitler's invasion of Paris, the inhabitants of Paris, Texas rose up in revolt. What follows is my story of how these people could be so mistaken, and what they need to realize to come to armistice. This is one case in which peace talks are all that is necessary; talk can actually settle the matter, if the combatants can be made to listen.

There seems to be no sharp date when the science wars began, but this much is clear: when anthropologists and sociologists, encouraged by works like Thomas Kuhn's *Structure of Scientific Revolutions*,[1] realized that scientists have a culture to

[1] Thomas S. Kuhn, *The Structure of Scientific Revolutions*, 3rd ed. (Chicago, IL: University of Chicago Press, 1996).

which sociological and anthropological methods apply, and began to apply them, there was a panicked reaction. People like Paul R. Gross and Norman Levitt, authors of *Higher Superstition*,[2] began to describe the new study of science as an attack on science. Some 15 years ago, the hostilities reached a boiling point with the infamous Sokal affair. Alan Sokal, a practicing physicist, decided to embarrass the enemy by infiltrating their ranks. He wrote an essay that was, as physics, manifest nonsense, but espoused many of the ideas of his ideological opponents in science studies. He then submitted it for inclusion in an interdisciplinary journal, *Social Text,*[3] to demonstrate the intellectual bankruptcy of the field of Science Studies. The journal accepted his paper for publication, at which point he announced his imposture, and the Science Wars erupted in frenzy. Naturally, those who wished to study science as a social institution tried to answer these charges, and so the fight was joined. Tensions have died down somewhat, but not because any resolution was reached.

The two sides of the conflict can be generally characterized this way: Some, especially scientists themselves, claim that science is, at least in the long run, immune to bias and self-correcting; others claim that science is a tool of political power, and therefore has biases built into it in favor of the powerful. Consider the following passages extolling the unbiased nature of science, one from philosopher of science Richard Rudner, one from scientist Carl Sagan:

> The books, so to speak, are never closed on any hypothesis in the precise sense that evidence relevant to the confirmation or disconfirmation of it can never be exhausted. Accordingly, it is fair to say that if a hypothesis we accept is false, the continued application of scientific method to its investigation will increase the likelihood that we will be able to correct our error by coming upon evidence that disconfirms it. It is in this sense, of a systematically built-in mechanism of corrigibility, that the intellectual history of the species has presented man with no more reliable ... method of inquiry than that of science.[4]

> Scientists, like all human beings, have their hopes and fears, their passions and despondencies – and their strong emotions may sometimes interrupt the course of clear thinking and sound practice. But science is also self-correcting. The most fundamental axioms and conclusions may be challenged. The prevailing hypotheses must survive confrontation with observation. Appeals to authority are

[2] Paul R. Gross and Norman Levitt, *Higher Superstition: the Academic Left and its Quarrels with Science* (Baltimore: Johns Hopkins University Press, 1994).

[3] Alan D. Sokal, "Transgressing the Boundaries: Towards a Transformative Hermeneutics of Quantum Gravity," *Social Text* 46/47 (1996): 217-252.

[4] Richard S. Rudner, *Philosophy of Social Science* (Englewood Cliffs, NJ: Prentice-Hall, 1966), 77.

> impermissible. The steps in a reasoned argument must be set out for all to see. Experiments must be reproducible.[5]

Similar expressions of confidence in the objectivity of science can be found in the writings of many scientists and philosophers of science. If science is guided by its very structure toward truth, then it cannot also be guided toward some other end (political domination), unless those ends coincide. I take it as obvious that they do not, or at least they usually do not; the set of propositions acceptance of which is useful to the powerful is not the same as the set of true propositions, or even the set of propositions most likely to be true given what else we now know. Consider now these passages, arguing for the politically biased nature of science, one from Marxist scholar Louis Althusser, one from philosopher of science Helen Longino:

> There is no perfectly transparent science, which, throughout its history as a science will always be preserved ... from ... the ideologies that besiege it; we know that a 'pure' science only exists on condition that it continually frees itself from the ideology which occupies it, haunts it or lies in wait for it. The inevitable price of this purification and liberation is a continued struggle against ideology itself.[6]

> The contextualist approach also makes it clear that the question of whether social values can play a positive role in the sciences is really the wrong question. Social and contextual values do play a role, and whether it is positive or negative depends on our orientation to the particular values in question. The feminist scientist, or the radical scientist, cannot simply try to be sensitive to the politically noxious values embedded in some research programs or try to avoid ideology by sticking to the data. 'Letting the data suggest' is a recipe for replicating the mainstream values and ideology that feminist and radical scientists reject.[7]

In other words, science is conditioned by ideology, and so does not correct itself with respect to that ideology. Similar expressions of doubt about the objectivity of science can be found in Bruno Latour,[8] Sandra Harding,[9] and a host of others.

These claims cannot both be true, and yet both seem plausible. It does seem that the scientific method (insofar as there is a single method) is designed precisely to root out error and tend toward truer and truer pictures of the world. On the

---

[5] Carl Sagan, *Broca's Brain* (New York: Ballantine, 1980), 96.

[6] Louis Althusser, *For Marx* (London: Allen Lane, 1969), 170.

[7] Helen Longino, *Science as Social Knowledge* (Princeton: Princeton University Press, 1990), 218.

[8] See for example, Bruno Latour, *Science in action: How to follow scientists and engineers through society* (Cambridge, MA: Harvard University Press, 1987), and Bruno Latour and Steve Woolgar, *Laboratory life: The construction of scientific facts* (Princeton: Princeton University Press, 1986).

[9] For example, Sandra Harding, *The Science Question in Feminism* (Ithaca, N.Y.: Cornell University Press, 1986), and *Whose science? Whose knowledge?* (Ithaca, N.Y.: Cornell University Press, 1991).

other hand, scientists are people, and scientific investigation is done by people in societies, and it would be amazing if they didn't bring their biases into the laboratory with them. We have two choices: we can take one side or the other, or we can find a way to reconcile the two views. I will argue that the two views are indeed compatible. When the proponents of the self-correcting nature of science say "Science is objective" and the proponents of science as an ideologically driven enterprise say, "Science is biased," they are not disagreeing, because they are talking at cross-purposes; they mean different things by the word 'science.' The former are talking about a method employed in theory choice, abstractly conceived; the latter are talking about a socially instantiated practice that has theory choice as a component. Consequently, it is possible for the abstractly characterized method of theory selection to be self-correcting, and yet be embedded in a larger practice that to some extent undermines, or even defeats, self-correction. In the end, the two sides can each have what they want.

This distinction between science as theory-choice procedure and science as social practice is easily confused with another, related distinction. For example, many scientists would admit that particular scientists may have let bias creep into their work, but that when they were doing so, they were doing bad science. In other words, it is ideal science, or good science, that corrects itself. But both sides to the debate can agree that there is good and bad science. The believers in science may admit that some scientists are biased, but they want to assert that it is not merely in the ideal that science corrects itself, but also in real practice. They want to claim that science as we actually do it has a tendency toward truth, which would be unwarranted if it were only science as ideally practiced that has that feature. Also, many of the political critics of science want to claim that even when science approaches the ideal of objectivity, it still serves political power. So the distinction between real and ideal science does not illuminate the problem.

Rudner, Sagan, and Sokal, along with the other defenders of the objectivity and self-correcting nature of science think of science as a method, structurally designed to weed out error. In particular, it is meant to weed out error due to the personal perspectives of scientists. The scientific method, as described in innumerable science textbooks, is something like this: a hypothesis is conceived, it doesn't matter how; logical consequences of that hypothesis are deduced; experiments are designed to see if those consequences are true; if not, the hypothesis is proven wrong, and the process returns to the beginning, with a revised or completely new hypothesis. If the consequences are correctly deduced, and the experiments are well-designed and well-performed, then the original hypothesis is refuted, even if it was the pet hypothesis of a well-beloved and authoritative scientist. Richard Feynman describes the method this way:

> In general we look for a new law [of physics] by the following process. First we guess it. Then we compute the consequences of that guess to see what would be implied if this law that we guessed is right. Then we compare the result of the computation to nature, with experiment or experience, compare it directly with observation, to see if it works. If it disagrees with experiment it is wrong. In that simple statement is the key to science. It does not make any difference how beautiful your guess is. It does not make any difference how smart you are, who made the guess, or what his name is – if it disagrees with experiment it is wrong. That is all there is to it. It is true that one has to check a little to make sure that it is wrong, because whoever did the experiment may have reported incorrectly, or there may have been some feature in the experiment that was not noticed, some dirt or something; or the man who computed the consequences, even though it may have been the one who made the guesses, could have made some mistake in the analysis.[10]

Feynman goes on to say that this picture is a bit oversimplified, but his further remarks only serve to add details to the three-part structure: hypothesis, deduction, experiment (see Figure 1).

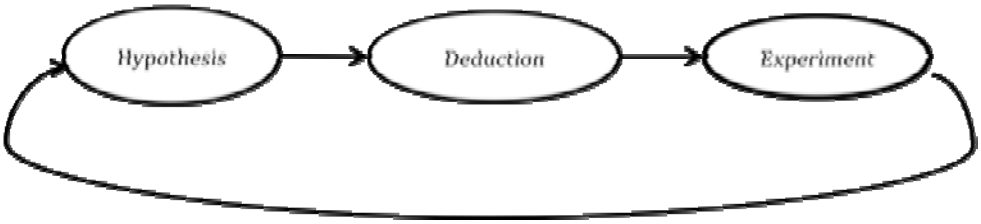

Figure 1

The results of the experiment then have an effect on what hypotheses get proposed, so the process is a self-correcting spiral, homing in on accurate representation of the world. It is easy to see how this understanding of what science is would lead one to think that it could not possibly be biased. If a biased scientist presents a faulty hypothesis, it will not be borne out by experiment, and so bias is rooted out, at least in the long run, by the structure of science itself.

II.

The critics of science as we now practice it do not see science as this idealized and highly abstract method of theory choice. The classical 'scientific method' is a component of science, but it is not the whole thing. They are thinking of science as a social practice that starts well before hypothesis with background information, distribution of resources and opportunities, and ends with publication and

[10] Richard Feynman, *The Character of Physical Law* (Boston: MIT Press, 1990), 156.

discussion of theories. What theories are accepted, published, and discussed forms the new background information out of which new hypotheses arise, so on this picture, too, science spirals, but the spiral is guided by more than just observation and experiment. It is because of these additional forces on scientific inquiry that science (in the 'practice' sense) can be biased, even if science (in the 'method' sense) is immune to bias.

Science as a social practice can be broken down into three stages: hypothesis selection, theory choice, and theory uptake. Theory choice has been the focus of much discussion of science, and so has become science itself for so many people, because it is amenable to abstract treatment. In particular, it is amenable to a normative understanding; understanding science as theory selection allows us to develop logics of science, and interpret particular cases of theory selection in terms of how well they achieve the goals of science, including an accurate picture of the world. But obviously there is more to how science gets done, and more to what scientific theories we accept, than the logic of theory choice alone. The scientific practice, as actually undertaken by real, working scientists, is better represented as a three stage structure, with theory choice taking place in a context of hypothesis selection and public uptake (see Figure 2).

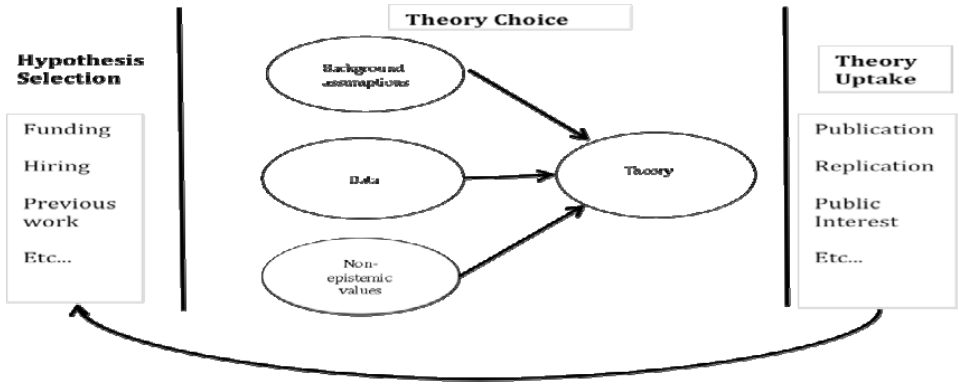

Figure 2

At the stage of hypothesis selection science gets its direction. To begin with, what science gets done is partly a function of what previous scientists have already done and what presently employed scientists would like to see done. Scientists are hired, promoted, and otherwise evaluated partly on the strength of how interesting the problems are that they are pursuing, so what we find out about the world is in part a function of what presently employed scientists find interesting. Hypotheses that no-one respects will have trouble finding funding and support; hypotheses that

are very radical will be difficult even to formulate, for lack of a history. So, what theories we accept is constrained by what hypotheses get tested. At the theory uptake stage there are similar constraints. If no scientific society or journal finds your work important or interesting, it will not get published, and so other scientists will not try to replicate the results, and the general public will never find out about it. Evolutionary biology had to wait decades for Gregor Mendel's groundbreaking work because it languished in a second-rate journal that nobody was reading. Even if a paper on a problem considered marginal by the majority makes it to publication, if the scientific community doesn't pick up on it, discuss it, and expand on it, it vanishes into obscurity. So when we confine ourselves to consideration of the scientific method, it is true that any hypothesis, no matter what it is or who brings it up, is treated equally; however, when we turn to the social practice of science, we see that only hypotheses that can attract enough interest to get resources, publication, and discussion really have a chance to be accepted.

To show that bias is possible, due to the social nature of science, is not to show that it is biased in favor of the powerful. That requires more argument. But that further argument is readily available. To put it simply, if our community has a political preference, it can and will express it in what science it supports with funding for research. The federal government, in the hands of conservatives, is much less likely to spend money on research in drug treatment as an alternative to punishment, studies of the harmful effects of religion on communities, fetal tissue research, birth control device trials, and the like. If our society considers the ailments of women uninteresting, then scientists who insist on studying women's ailments will lose funding, respect, promotions, and employment opportunities. They will also find it hard to publish their results, no matter how well-designed their studies are. In all these ways, the communities of which we are part can help some projects and hinder others, all because we need those communities' help to do our research. If science has been biased in the past, then the research projects against which it is biased will have less of a history. In other words, there will be less previous work to draw on, so the work will be correspondingly more difficult. Research tools, like indices, can also be skewed against some research projects. If there is no category in *Psychological Abstracts* for Abused Woman Syndrome, then people who want to do research on that have a harder job tracking down what previous work there is.

The defenders of science might at this point concede that science-the-social-practice can be and probably is politically biased, but that the bias is safely confined to the stages before and after theory-selection. Consequently, they may affect the direction that science takes, but the structure of the scientific method will insure

that, at least in the long run, we will be protected from accepting false theories. So we may overlook some interesting avenues of research, but we will not accept bad theories. There's something right about this. There are two kinds of bias, from two parts of inquiry: directional bias, that is, bias in what problems science pursues and what problems it ignores, and experimental bias, which is bias in how theories are chosen on the basis of the scientific method. Directional bias by itself will not lead us into error; the most it will do is lead us to miss some of the truth. Experimental bias is the kind of bias that causes trouble, because it will lead us to accept false theories. But the scientific method is supposed to correct for that, so we are only in danger insofar as we do science badly. Careful attention to experimental design and interpretation of results will at least keep us moving toward truer and truer theories. Unfortunately, this is not quite right. The picture of the scientific method schematized in Figure 1 is inaccurate, even as an idealization.

III.

Bias causes two different kinds of undesirable result: error and ignorance. In other words, bias can lead you to accept something false because you find it congenial, or it may lead you to ignore some truths because you find them uncongenial. To say that science is self-correcting and therefore structurally immune (in the long run) to bias may mean only that it is protected from error in the long run, not from ignorance. Experimental bias leads to error, directional bias to ignorance. It seems that the critics of science are right about the biases that occur in hypothesis selection and theory uptake. The defenders of science as an abstract method are right to point out that the scientific method is particularly well designed for avoiding error, but this does not mean that they are right about the immunity of the method from error, even in the long run. As nice as it would be to say that both sides are right, each in its own way, once the ambiguity of the term 'science' is pointed out, the fact is that even theory selection abstractly conceived is not immune to bias. The picture of the scientific method schematized in figure 1 is oversimplified.

Theories are supported by more than mere observation. Any set of observations is consistent with an infinite variety of different theories. Consequently, when we decide between empirically equivalent theories, we are always deciding on the basis of more than observational data. One way we narrow down the number of contender theories is by invoking nonepistemic values, like simplicity, predictive fruitfulness, and elegance. When scientists are faced with mutually inconsistent but equally empirically adequate theories, they frequently choose which to pursue in further research on grounds other than likelihood of truth. It may be, as Willard Quine (in some moods) says, that two theories that are completely empirically

equivalent – that is, entail all the same observations – are mere verbal variants on one another. But we are not at liberty to say that about theories that are equivalent only in that they entail the same evaluations for observation sentences we already know the truth-values of, or ones we are likely to know soon. This limited kind of empirical equivalence is the pressing problem, and equally intractable if we limit ourselves to differences between the theories that bear on how likely they are to be true. Similarly, and also because of the underdetermination of theory by data, scientists judge particular theories in the light of what they already believe. If a scientist earlier in this century had a firm belief that all the laws of nature must be deterministic in character, not statistical, then she would be inclined to reject the standard understandings of quantum theory and accept hidden-variable versions of the theory. Other scientists accept the standard understandings of quantum theory as superior to hidden-variable theories (on the grounds of simplicity) and therefore reject the view that all the laws of nature must be deterministic. It is probably right to reject the deterministic view of law rather than the value of simplicity, but the point is, whether any given instance is right or wrong, these kinds of considerations have a role to play in theory choice. This feature of science makes it possible for political biases to play a role even in theory selection. If our society considers one kind of person inferior to another kind of person, then scientists will bring that prejudice as an auxiliary assumption to their research. They will tend to accept theories that support that claim, in spite of flaws, and be skeptical of theories that challenge it, in spite of support by data.[11]

Another way we narrow down candidate theories is by appeal to non-epistemic theoretical values. It is hard to see how what values we use to judge theories could be employed to political ends. Nevertheless, it almost certainly happens that theories are chosen over other theories for reasons that have nothing to do with whether they are likely to be true, and people with different political aims would accept different theories. Suppose for example that the two theories under consideration are indistinguishable with respect to empirical consequences, simplicity, and whatever other uncontroversial theoretical virtue you like. If the first theory is true, then a commercial application that will lead to gigantic profits will follow. If the second theory is true, then there is no readily exploitable commercial result, but there is an easy application to some problem in the third

[11] For two examples of cases in which this sort of thing seems to have happened, see Nancy Tuana's "The Weaker Seed: The Sexist Bias of Reproductive Theory," in her edited book *Feminism and Science* (Bloomington: Indiana University Press, 1989), and Stephen Jay Gould, *The Mismeasure of Man* (New York: Norton and Co., 1981). Longino, *Science as Social Knowledge,* is also full of examples.

world. People in science for the sake of technological application and profit will tend to accept and pursue the first theory, while people who see the role of science as one of improving quality of life for people will tend to accept the second theory. The theoretical value in this kind of case is tied to what the scientist takes to be the primary role of science in society, what 'good science' is supposed to do. Of course, if those theories really are alternatives, then one or both will be wrong, but it may take quite a bit of time to find out which, and in the meantime, science will continue in the direction it has taken because of a non-epistemic value choice. The upshot of all this is that science, because of its essentially social nature, can be and frequently is bent to political purposes, in spite of its tendency to root out certain kinds of errors.

So the peace plan is simple. Those who stand with Rudner, Sagan, Feynman, and Sokal can have their territory in logical space – that the scientific method, if applied correctly and with honest intent, will tend to expose error – as long as they have no further territorial ambitions, for example, to the claim that this gives scientists special social authority. Those who stand with Latour, Longino, Althusser, and Harding can have their territory – that scientists are people in a social setting, susceptible to political pressures that can and often do invisibly affect how their results come out – provided they have no further territorial ambitions, say, to the claim that there is no objective fact to which scientific theory aims to approximate. It may well be that scientists are especially objective folk and should have a special social authority, but the mere utility of the science-as-method doesn't show that. It may be that reality is radically socially constituted, but the mere social nature of science-as-practice doesn't show that.[12]

[12] An earlier version of this paper was delivered at the North Texas Philosophical Association meeting in April of 1997, and at the conference *Webs of Discourse: the Intertextuality of Science Studies*, held at Texas Tech University in February, 1998.

DEBATE

ON EPISTEMIC ABSTEMIOUSNESS: A REPLY TO BUNDY

Scott F. AIKIN, Michael HARBOUR,

Jonathan NEUFELD, Robert B. TALISSE

ABSTRACT: In the previous issue of *Logos & Episteme*, Alex Bundy defends epistemic abstemiousness and criticizes our tales of epistemic abstention to martyrdom and conversion. Here we will briefly reply to Bundy's criticisms and then diagnose what we think the trouble is with the going versions of epistemic abstemiousness.

KEYWORDS: epistemic abstemiousness, epistemic martyrdom, epistemic conversion, Alex Bundy

Many thanks to Alex Bundy for his defense of epistemic abstemiousness in the previous issue of *Logos and Episteme*[1] and to the editors for the opportunity to reply to his essay. Bundy criticizes our tales of epistemic abstention to martyrdom and conversion along two lines. First, he argues that the consequences of conversion, as we tell the story, are not reasons that count against epistemic abstemiousness. Second, Bundy argues that in cases of dogmatic behavior from an interlocutor, we have evidence against that speaker's peerhood. Here we will briefly reply to Bundy's criticisms and then diagnose what we think the trouble is with the going versions of epistemic abstemiousness.

The main target for criticism in our original essay[2] was what we called the *principle of suspension* (PS), which Bundy nicely indexes for our purposes: *if S disagrees with a peer about* p*, and S is aware of the disagreement and of the peer as a peer, then S should suspend judgment about* p*.* As we said earlier, we have a good deal of sympathy for this principle and the broader program of epistemic abstemiousness. It is not only *prima facie* plausible, but seems the exact sort of responsible line to take in the face of deep disagreement between peers. However, our story shows that deploying the principle has untoward consequences. Consider the consequences of what we've called epistemic martyrdom. Betty and Alf are

[1] Alex Bundy, "In Defense of Epistemic Abstemiousness," *Logos and Episteme* II, 2 (2011): 287-292.

[2] Scott Aikin, Michael Harbour, Jonathan Neufeld, and Robert B. Talisse, "Epistemic Abstainers, Epistemic Martyrs, and Epistemic Converts," *Logos and Episteme* 1, 2 (2010): 211-9.

peers and disagree about p. Betty, upon discovering this disagreement, follows PS, and she thereby puts Alf in (it seems at least to him) a better epistemic position by suspending judgment. Alf now has no one directly disagreeing with him, and so has had a potential defeater simply vanish. This seems strange that one subject following PS should improve the dialectical situation for another opposing subject.

Bundy objects that "norms like PS are not evaluated according to how they contribute to the epistemic good of some other person or sets of people."[3] That is, PS is a rule for an *individual*, not for *groups*, and so Bundy's argument is that if the epistemic virtues of individuals yield group irrationality, then so be it. This seems a curious bullet to bite to save abstemiousness, as these principles are ones that are deeply other-regarding – they are about disagreement and taking the views of opponents very seriously. Those who care for PS and its abstemious kin are motivated by a care for what and how their peers think, and so it is a troubling way to save those principles by saying that the cognitive consequences for those peers do not matter to PS's evaluation.

Bundy's second line of criticism is of our account of the path from epistemic martyrdom to epistemic conversion. Our tale was that once Betty has suspended regarding p, Alf's assessment of the situation has improved his commitment to not-p. Betty now has new meta-evidence, because the dialectical situation regarding p has changed. Moreover, because Betty is committed to PS, it seems she must adopt a similarly abstemious commitment to its application in this circumstance – if Betty disagrees with Alf about the proper response to the dialectical logjam, then she (by PS) should suspend about whether she should suspend judgment regarding p. And when that is the case, she is returned again to Alf's confident acceptance as meta-evidence. Bundy thinks this is 'double-counting' on Alf's part (and by extension, Betty's), but aren't these different pieces of evidence? They occur at different times and they are responses to different dialectical circumstances – the first regarding p, the second regarding p and whether PS applies to this case. Take a similar sort of circumstance, perhaps the question of whether a joke is funny. First, you may tell the joke, and the audience does not laugh. That is evidence your joke is not funny. You could then remind the audience of certain principles of humor and explain the joke to them. They still don't laugh. The fact that the audience persists in not laughing is not simply the same evidence as their initial reaction that the joke isn't funny, but is more evidence. The same goes for Alf and Betty. Bundy would have it that a peer's doxastic attitude is transparent to the reason that caused her to hold it, so that, once we see the reason, we no longer take the attitude as evidence. But this

[3] Bundy, "In Defense," 289.

is to deny the central point of PS: the doxastic attitudes of our peers matter to our deliberations.

At this point, though, Bundy thinks that Alf has shown himself (at least to Betty) to be Betty's epistemic inferior. He holds, "Alf is no longer a peer when it comes to p."[4] As a consequence, Bundy takes it that the untoward series of events from Betty's suspension to martyrdom to conversion needn't unfold, because a Betty with some judgment wouldn't go down such a path. Crucial to Bundy's objection, though, is that Betty reason along the following lines:

> [N]ow Betty has good evidence that Alf is not living up to his epistemic duties – namely, he is not conforming to PS! So after the first adjustment of her belief regarding p, Betty does not have to continue to apply PS until she ends up believing not-P.[5]

So, according to Bundy, Betty may use the PS as a rough criterion for peerhood. This strategy would work were PS not self-reflexive (i.e., were PS to run that when S knowingly disagrees with a peer regarding p, S must suspend judgment with regard to p, except for when p=PS). The trouble is that there's no obvious reason why the PS can be made so as to except itself. If *prima facie* Alf and Betty are peers and their disagreement about p does not undermine that, then why does their disagreement about PS count as a case for undermining peerhood? It would surely be strange to say: *I thought we were roughly intellectual peers until I found out all the things we disagree about, especially epistemology*. Even if it were possible to exempt PS from peer disagreement, the consequences seem to be against the spirit of the principle.

Bundy doesn't dispute that some of the deepest and most important questions are likely to spark just the kind of disagreement that we point to in our paper – often sparking second order disagreements. In cases involving disagreement over the application of PS, though, the both first- and second-order disagreements are simply short-circuited by denying peerhood to Alf. So while it seems that there is deep disagreement among peers about matters that are important to us, we should take heart: we have far fewer peers than we originally thought.

Indeed, on Bundy's view, it turns out that epistemic disagreement amongst peers is *impossible*. This is because, in order to be genuine peers, two parties must be capable of appropriately applying PS. So whenever two parties are locked in a disagreement over *p*, they will, if they are in fact peers, *agree* to suspend judgment with respect to *p*. This is an odd result. *Epistemic abstemiousness* is motivated by the fact that disagreement amongst apparent peers is a basic feature of our epistemic

[4] Bundy, "In Defense," 291.

[5] Bundy, "In Defense," 291-2.

condition; even after extensive deliberation, people often still disagree. But if Bundy is right, this problem is illusory. Deliberation amongst peers can never terminate in disagreement because *true peers* will always resolve disagreement by suspending judgment.

We should be clear that we are sympathetic with the PS and the broader program of epistemic abstemiousness. Our trouble is that the principles of abstention, as stated, are only *synchronic*, and they may be right from the perspective of time-slices of cognitive duties, but they have *diachronic* consequences, too. One may, *looking at one time*, be doing what's right, but *over time*, it could be wrong. For example, take the hedonist's rule of maximizing pleasure and minimizing pain. Surely without a diachronic view, following the rule would ultimately lead to its frustration. One must take on long-term goals, plan, make sacrifices, and so on. The same, we think, goes for epistemic principles that have practical bearing, because when one follows an other-regarding epistemic principle, especially in a dialogue, one thereby changes the situation the principle bears on.

RECOVERING RESPONSIBILITY

Guy AXTELL

ABSTRACT: This paper defends the epistemological importance of 'diachronic' or cross-temporal evaluation of epistemic agents against an interesting dilemma posed for this view in Trent Dougherty's recent paper "Reducing Responsibility." This is primarily a debate between evidentialists and character epistemologists, and key issues of contention that the paper treats include the divergent functions of synchronic and diachronic (longitudinal) evaluations of agents and their beliefs, the nature and sources of epistemic normativity, and the advantages versus the costs of the evidentialists' reductionism about sources of epistemic normativity.

KEYWORDS: epistemic normativity, character epistemology, evidentialism, synchronic and diachronic

1. Epistemic Normativity and the Synchronic/Diachronic Divide

Trent Dougherty's recent article, "Reducing Responsibility: An Evidentialist Account of Epistemic Blame,"[1] raises quite interesting issues about the epistemic appraisal of agents and their beliefs. Dougherty, editor of a new collection, *Evidentialism and its Discontents*,[2] constructs and discusses a case similar to those that authors he terms the 'core responsibilists' have utilized as counter-examples to evidentialism. But Dougherty's purpose in constructing his case – the 'Craig Case' as we'll call it – is to defend the analysis of epistemic justification advocated by evidentialist thinkers like Earl Conee and Richard Feldman, according to which 'evidential fit' (or 'synchronic epistemic rationality') is the sole source of properly epistemic norms. This he terms the evidentialists' 'reductionist' account of epistemic normativity, in sharp relief from the 'non-reductionism' of the core responsibilists, who allow and indeed insist upon other sources of epistemic normativity besides evidential fit.

As technical background, Conee and Feldman hold that a proposition being justified for an agent is just a question of that agent having at any and every given moment of time that singular doxastic attitude towards that proposition – belief,

[1] Trent Dougherty, "Reducing Responsibility: An Evidentialist Account of Epistemic Blame," *European Journal of Philosophy*, 2011, doi: 10.1111/j.1468-0378.2010.00422.x (accessed August 2011).

[2] Trent Dougherty, ed., *Evidentialism and its Discontents* (Oxford: Oxford University Press, forthcoming 2011).

suspension, or disbelief – that uniquely 'fits' the total evidence bearing upon it that she possesses. Propositional justification is thus a matter of the right sort of cognitive response to evidence (mentally-accessible reasons), and this is wholly a 'synchronic' (present time-slice) evaluation. The evidentialists furthermore holds the propositionalist view that epistemic justification (justification that epistemizes true belief, i.e., renders it knowledge) is just the having of propositional justification for some proposition, plus the basing of one's actual belief upon those reasons.[3] Their view of 'epistemic normativity' derives directly this propositionalist account, although it is a contentious thesis not shared by non-internalists, who typically begin with a conception of the basing relationship and of *doxastic* justification that does *not* entail propositional justification as a necessary condition on knowing.

The case Dougherty has us focus on is that of an imaginary friend and peer, Craig, a short earth creationist who basically ignores his (Dougherty's) introduction of counter-evidence to this belief as well as a well-meaning recommendation that Craig consider and investigate this counter-evidence by reading a book he offers him on subjects of geological and evolutionary science. The worry in the Craig case and in many cases like it is that the agent *is and remains* personally justified according to the evidentialist standard of epistemic fit, but ironically only because that agent's evidence-base is so extremely narrow; indeed, Craig nurtures this ignorance of alternatives by simply dismissing or otherwise failing to pursue inquiry into this potentially undermining counter-evidence to his belief. "[E]videntialism requires that a change in epistemic status of belief issue from a change in evidential status."[4] But the Craig Case is one where the personal justification of the agent appears to drop with or after the encounter with his friend, and for reasons *other than* lack of evidential fit.[5] A responsibilist at first gloss might suggest that the drop

[3] Jonathan Kvanvig, "Propositionalism and the Metaphysics of Experience," *Philosophical Issues* 17 (2007): 165-78.

[4] Dougherty, "Reducing Responsibility," 5.

[5] I insert 'drop' along with simple 'loss' of status in recognition of the important point that externalists and holders of 'mixed' accounts beleive that Craig *never* had the positive status of being epistemically justified (doxastically justified) in the first place, by virtue of meeting the evidentialist standards of doxastic justification (having evidence that propositionally justifies, plus basing one's belief on that evidence). Where diachronic justification (competent or responsible evidence-gathering) isn't a further requirement, agent reliability is by no means ensured, and doxastic justification or warrant as externalists understand it, won't be present. Yet, as Baehr writes, "Perhaps there is some epistemic value simply in having a belief that fits one's evidence – regardless of whether this evidence is the result of defective inquiry. Such beliefs might be said to involve a kind of logical coherence or consistency, which indeed is often regarded as an epistemic desideratum." (Jason Baehr, "Evidentialism, Vice, and Virtue," *Philosophy and Phenomenological Research*, 78 (2009): 549.)

in personal justification for his belief is due to Craig subsequently manifesting poor intellectual motivations (strong confirmation bias) and his displaying intellectual irresponsibility diachronically (cross-temporally) through failing to engage or 'deal with' the counter-evidence to his belief in ways that we would reasonably demand of any intellectually virtuous agent similarly situated.

Dougherty readily agrees that the evidentialist must take account of such cases. However, he argues, on closer inspection evidentialism is well-able to handle them by offering a kind of *error theory*. "My position is that all instances of epistemic irresponsibility are in fact either forms of instrumental rationality or moral irresponsibility in so far as there is anything amiss that goes *beyond* one's belief fitting the evidence one has at the time."[6] Despite first appearances, then, there is nothing distinctly epistemic about what the responsibilists call 'intellectual responsibility' in inquiry: being diachronic or longitudinal evaluations, these are misconstrued "moral and prudential evaluations of behavior related to the formation of beliefs."[7] To show this Dougherty offers three different responses which an evidentialist might make, which he thinks don't require us to recognize anything distinctively epistemic about Craig's failures and which therefore don't require us to recognize any exception to the evidentialist account of epistemic normativity.

At issue between the evidentialists Dougherty defends and the self-described responsibilists he targets with his error theory is not only the role of time-slice (synchronic) and longitudinal (diachronic) evaluations in epistemology, but also, equally important to him, the right attitude that a philosopher should hold towards reductionism about the sources of epistemic normativity. The intended upshot of his treatment of the Craig Case is to re-instate the evidentialists' reductionist account of epistemic value by replying directly to the best responsibilist counter-examples. His reductionist thesis as localized to the responsibilist's person-level or diachronic picture of epistemic evaluation is formalized as the Identity Thesis:

> IT: Each instance of epistemic irresponsibility is just an instance of purely nonepistemic irresponsibility /irrationality (either moral or instrumental).[8]

But this is just one application of a more general reductionist approach to *epistemic normativity*, captured in the statements that "any *normativity* concerning

---

[6] Dougherty, "Reducing Responsibility," 3. "Like Feldman, my position is that when one's belief fits the evidence, all other forms of evaluation concerning the belief are either moral or instrumental." (Dougherty, "Reducing Responsibility," 5.)

[7] Dougherty, "Reducing Responsibility," 3. Compare Richard Feldman, "The Ethics of Belief," in *Evidentialism: Essays in Epistemology*, ed. Earl Conee and Richard Feldman (Oxford: Oxford University Press, 2004), 190.

[8] Dougherty, "Reducing Responsibility," 4.

belief that goes beyond fitting the evidence, and in particular epistemic responsibility, is either moral or instrumental"[9], and that "Like Feldman, my position is that when one's belief fits the evidence, *all other forms of evaluation concerning the belief* are either moral or instrumental."[10]

As one of the 'core responsibilist' authors targeted by Dougherty in his paper, I want to make my replies to this reductionist thesis and to his application of it in treatment of the Craig Case, while recognizing that I cannot speak for any other author treated in this paper. Other thinkers critical of the Identity Thesis might take a significantly different tact in their own responses. There are a number of interesting questions that arise in the above quotations, but I first want to point to two less obvious ones, questions that Dougherty's implicit answers to already inform the manner in which he glosses the concept of epistemic normativity in his paper. One of these questions is whether epistemic normativity applies only to evaluation concerning beliefs, or also to evaluations concerning agents and their habits and dispositions. As Roger Pouivet aptly notes, "Virtue epistemologists generally agree that, more than anything, good intellectual *habits* ground our pretensions to warranted beliefs, and to knowledge. And habits are properties of persons, not of beliefs."[11] Another is whether epistemic normativity, even granting it the narrower extension of '*normativity* concerning belief,' is best construed as restricted to knowledge-relevant epistemic status, or includes *theoretic understanding* among primary positive epistemic goods or standings. Responsibilists have been among the foremost proponents of conceiving the *telos* of the life of the mind to include (at least) understanding, and this question of epistemological axiology is directly relevant to the question as to whether diachronic assessments of agents can be properly epistemological. What Raymond Nickerson in *Aspects of Rationality* calls 'active fair-mindedness' – an interpretation that would require one to put "significant effort into seeking evidence, including evidence that goes against a favored hypothesis as well as evidence that supports it"[12] – may not seem to be required for more passive sorts of propositional acceptance, yet seems prerequisite for achieving the epistemic good of understanding. Virtue-relevant epistemic standings such as theoretical understanding are thus more naturally given to a dynamic and developmental account than is true belief, often conceived (or misconceived) as a static cognitive state. We cannot directly pursue these matters here, but we can

[9] Dougherty, "Reducing Responsibility," 1, emphasis added.

[10] Dougherty, "Reducing Responsibility," 5, emphasis added.

[11] Roger Pouivet, "Moral and Epistemic Virtues: A Thomistic and Analytical Perspective," *Forum Philosophicum: International Journal for Philosophy*, 15, 1 (2010).

[12] Raymond Nickerson, *Aspects of Rationality* (New York: Taylor and Francis Group, 2008), 140.

return to consider their place in this debate after we attend more directly to explicating Dougherty's thesis and argument. Put in more positive terms, the reductionism of the evidentialists that he argues in support of is that evidential fit is the only proper source of epistemic normativity, and that this is wholly a matter of synchronic evaluation.

Michael Williams' term 'due care and diligence,' from his paper "Responsibility and Reliability"[13] will be used here to characterize the responsibilist position on the properly epistemological standing of evaluations of an agent's motivations and habits that bear directly upon the conduct of inquiry.[14] Dougherty writes, "[T]he core responsibilists all say that there is more to epistemic responsibility than evidential fit" and that if evidentialism is false, then the ethics of belief goes well beyond this consideration of synchronic rationality.[15] This is quite correct; there *is* considerably more to epistemological evaluation – and still more especially to the ethics of belief – than is allowed for by a theory that takes believing according to one's present total evidence (however gotten or ill-gotten) as the only source and measure of epistemic value. I think the issues that his paper raises are misdescribed as issues about the ethics of belief, because that is a question of what one should believe 'all things considered,' which is not Dougherty's focus. The ethics of belief, taken this way, explicitly involves ethically-guided considerations of the consequences our beliefs have on our actions, and so obviously goes well beyond considerations of synchronic rationality or evidential (propositional) justification. (10) What Dougherty seems to have in mind is something closer to 'belief's own' ethics – doxastic norms. Having argued elsewhere against both sides of Feldman's evidentialist progam – his account of 'epistemic' justification since the driving idea is, and what he describes as ethics of belief[16] – I here only want to argue that there

[13] Michael Williams, "Responsibility and Reliability," *Philosophical Papers* 37, 1 (2008): 1-27.

[14] Williams distinguishes two senses of epistemic responsibility, as a matter of *accountability* (responsibility *for* belief – we are responsible for what we believe in the sense of *accountable*) and as it contrasts with *ir*responsibility: failure to exercise due care or diligence in what I term our inquiry-directed or *zetetic* habits and strategies (sometimes as attested-to by our overt activities or behaviors). The antonym of the first sense is *not responsible,* and of the second sense is *irresponsible.* Williams writes that "The two kinds of responsibility are related in an obvious way: We are accountable for our beliefs precisely because of our obligation to manage them properly." (Williams, "Responsibility and Reliability," 2.) The issue between Dougherty and the responsibilists is primarily over this second sense. The responsibilists hold that knowledge involves reliability, and that justified belief is justified (responsible) *believing* (i.e., doxastic justification or warrant).

[15] Dougherty, "Reducing Responsibility," 10.

[16] Guy Axtell, "From Internalist Evidentialism to Virtue Responsibilism," in *Evidentialism and its Discontents*, ed. Trent Dougherty (Oxford: Oxford University Press, forthcoming).

is more to 'epistemic normativity' and to a proper understanding of responsibility-relevant norms than evidential fit and the agent's having the sort of rationality that it is characterized by. The agent with unconscientiously acquired or maintained true belief is not necessarily doing fine qua epistemic agent, even when their ascent to a proposition is finely in line with their evidence. Unconscientiously held beliefs will definitely affect one's understanding, but I think there is more to it even if this term 'epistemic normativity' is restricted just to 'knowledge-relevant' normativity. Evaluation of *persons* (agents), as contrasted with the evaluation only of beliefs (cognitive states), is properly part of the theory of knowledge, and evaluation of persons is never wholly a synchronic affair abstractable from motivational factors and from habits of inquiry. In short then, my stance is that epistemic fit is not a sufficient or even a very useful measure of personal justification, epistemic value, or agent reliability when set apart, as evidentialists intend, from *diachronic* considerations of how well or poorly motivated and conducted were the inquiries that provided the agent with just *that* set of what the evidentialist calls an agent's 'total evidence at time t.'[17]

Dougherty to his credit provides very clear characterizations of the issues and of the stances various philosophers have taken, and one thing I agree with fully in his paper is that for the responsibilists' explicitly non-reductionist account to be plausible, its proponents need to show that there are nontrivial cases of what they call epistemic irresponsibility that cannot be better explained in one of the three explanations he offers in the Craig Case. That case presents a common (all-too-common) situation, and (with one notable exception to be discussed below) is as good as most cases might be in serving as a focal point for this debate. The thrust of my reply is therefore a straightforward one: None of the alternative explanations Dougherty offers in the Craig Case is as plausible as he thinks they are; neither individually nor collectively do they provide evaluations *as* plausible or informative as the one according to which the personal justification Craig enjoyed for his creationist belief does indeed drop, and drops not merely due to the discovered *fact* of peer disagreement, but due to Craig's own subsequent failure to meet an expected (diachronic) standard of due care and diligence in treating the counter-evidence his friend and peer presents him with.

[17] The virtue responsibilists can certainly agree with their virtue reliabilist counterparts that the etiology of belief matters in knowledge-yielding doxastic justification, and in the achievement of other epistemic goods as like theoretical understanding as well. Synchronic rationality *qua* 'fit' of the strength of belief with one's evidence at any point in time is thus viewed as at best one among a number of sources of epistemic standards, which include agent reliability as well as reliability–enhancing intellectual motivations and habits of inquiry. (Guy Axtell, Philip Olson, "Three Independent Factors in Epistemology," *Contemporary Pragmatism* 6, 2 (2009): 89-109.)

2. Reductionism, Non-Reductionism, and the Unification Ideal

Before proceeding to the Craig Case, the three different possible evidentialist explanations of it, and my responses to the adequacy of each, something more should be said about Dougherty's contrast of reductionism and non-reductionism. He acknowledges that there are a range of positions that might support non-reductionism as he employs that term, some but not all being versions of virtue epistemology. He cites the more explicitly non-reductionist authors like Lorraine Code, James Montmarquet and I, among those who develop the analogies between ethical and epistemic evaluations. For the virtue responsibilists non-reductionism does not preclude, as one might initially think, but even *supports* aspirations to greater theoretical unity between philosophy's main normative sub-fields.[18] It is a possible path to properly circumspect theoretical unity, and as such not to be dismissed as an obstruction to it or a merely eclectic view. Dougherty rightly notices that the non-reductionist approaches of certain self-described responsibilists are considerably different than the approach taken by Linda Zagzebski, who in *Virtues of the Mind* tries to "subsume the intellectual virtues under the general category for moral virtues."[19] This Dougherty is right to point out is a subsumption thesis, and a reductionist thesis albeit one running in a direction quite *contrary* to the evidentialists' own intended reduction.

Yet we should draw attention to *more* positions available to non-reductionists than Dougherty distinguishes; Susan Haack[20] helpfully identifies five distinct views in the literature about the relationship between ethical an epistemic appraisal:

> [1] that ethical appraisal is strictly inapplicable where epistemological appraisal is relevant (the *independence thesis);*
>
> [2] that epistemic appraisal is distinct from, but analogous to, ethical appraisal (the *analogy thesis)*
>
> [3] that positive/negative epistemic appraisal is distinct from, but invariably associated with, positive/negative ethical appraisal *(correlation thesis);*
>
> [4] that there is not invariable correlation, but partial overlap, where positive/negative epistemic appraisal is associated with positive/negative ethical appraisal (the *overlap thesis)*

---

[18] Although similarly confused identifications of unification and reductionism do sometimes occur even in the sciences, most philosophers of science acknowledged that these are separable goals; it seems possible then to hold that a virtue theory or other general theory of value provides a non-trivial degree of unification to epistemology and ethics.

[19] Linda Zagzebski, *Virtues of the Mind* (Cambridge: Cambridge University Press, 1996), 255.

[20] Susan Haack, "The Ethics of Belief Reconsidered," in *Knowledge, Truth, and Duty*, ed. Matthias Steup (Oxford: Oxford University Press, 2001), 21-34.

> [5] that epistemic appraisal is a subspecies of ethical appraisal (the *special-case thesis).*

On this taxonomy, Dougherty holds Model [1], and Zagzebski's claim in *Virtues of the Mind* "that epistemic evaluation is a form of moral evaluation"[21] is an instance of Model [5]. Non-reductionists might hold any of Models [2]-[4], some of which it should be noted make considerably stronger claims than others; yet each of [2]-[4] are able to consistently maintain basic distinctions between theoretical and practical reason, as well as between epistemic and moral evaluation (as Model [1] arguably is not). This broadened taxonomy is helpful in correcting one mistake it appears Dougherty makes in characterizing the non-reductionism of the core responsibilists and the burden of proof issues that lie between them. For at one point he strays from his initial, correct characterization and attributes to them a stronger thesis than I think its defenders need to or in fact do make. This comes when he claims that they need to demonstrate "a purely epistemic category of evaluation which does not concern fit with one's evidence," or again, "a purely epistemic evaluation" or kind of irrationality, "over and above synchronic irrationality."[22] It is quite unclear what 'purely' is meant to indicate in these passages, and credit-worthy intellectual habits and problem-solving strategies are anyway surely connected not *only* with concerns of agent reliability in the external sense, but also with concerns of how evidence, including counter-evidence, is internally processed by agents engaged in inquiry. But Dougherty's claim about what the responsibilists *must demonstrate* in order to maintain their non-reductionist stance seems to function as a burden shifting move that they needn't accept. Even the strongest of the three models ([2]-[4]) available to them asserts *only* a strong kind of 'entanglement' of epistemic and ethical evaluation (Model [4], the *overlap thesis,* which happens to be Haack's own view).[23] Yet this strong kind of entanglement, and even the weaker kinds of it in Models [2] and [3], the *analogy*

[21] Zagzebski, *Virtues of the Mind*, 6.

[22] Dougherty, "Reducing Responsibility," 6.

[23] This should be familiar from the history of ethics in the 20th century, where for instance the position of pragmatists from Dewey through Putnam, in opposition to the influential schemes of reduction offered by C.L. Stevenson, and typically endorsed by the logical positivists, took the form of entanglement thesis (see Putnam's *The Collapse of the Fact/Value Dichotomy and Other Essays* (Cambridge, MA: Harvard University Press, 2004)). Defenders of the irreducibility of 'thick' affective and characterological concepts to 'thin' deontological ones is another pertinent example; in writers from Bernard Williams to John McDowell this irreducibility is similarly couched as an entanglement thesis (quite contrary to asserting a 'pure x' idea). See also Putnam, *The Collapse of the Fact/Value Dichotomy*, for an extended discussion of return of philosophical interest in thick concepts and its relationship with "the collapse of the fact/value dichotomy."

thesis and the *correlation* thesis, appear substantial enough to motivate certain main tenets of responsibilist thought, including especially the inclusion of 'character epistemology' (Baehr's term) into epistemology's purview and the properly epistemic standing of evaluations of intellectual motives and habits. This is why I thus think it is mistaken to claim that it is incumbent upon critics of the Identity Thesis to demonstrate "a purely epistemic category of evaluation which does not concern fit with one's evidence."[24] This is misguided because the responsibilist can hardly be expected to argue for entanglement (strong or weak) and for 'purity' at the same time. Better to acknowledge, as social and feminist epistemologists like Heidi Grasswick do, the "impurities of epistemic responsibility," and sharing their demand "that a viable concept of epistemic responsibility must be consistent with the impurities of epistemic agency … as we make decisions regarding how to know."[25] The only *direct* burden of a non-reductionist as Dougherty uses that term is to show that proposed reductions *don't work*. Yet the same point also underlines that it *is* clearly incumbent upon Feldman and other reductionists to demonstrate that every instance of purported epistemic irresponsibility in the kinds of cases before us is really "just an instance of purely non-epistemic irresponsibility/irrationality" – *since that is just what Identity claims.*

3. Competing Explanations in the Craig Case

With these notes about the Identity Thesis and the nature of the responsibilists' opposition to it out of the way, we can precede directly to Dougherty's case of Craig the short-earth special creationist. The author acknowledges finding Craig's mind-set troubling:

> The problem, though, didn't seem to be that his beliefs didn't fit his evidence—they did seem to fit his evidence, for he had read very narrowly on the subject and had been raised and schooled all his life in an apparently reliable community which sustained this belief in the usual social ways, and which had reasonable-sounding stories for why people deny their views. Rather, the problem seemed to be precisely that he only had the very limited evidence he had, since I'd often recommended books challenging his views. In language that is becoming more common, his belief seemed to satisfy the standards of synchronic rationality: it seemed to fit the evidence he had at the time; but it didn't appear to meet the

[24] Dougherty, "Reducing Responsibility," 5.

[25] Heidi E. Grasswick, "The Impurities of Epistemic Responsibility: Developing a Practice-Oriented Epistemology," in *Recognition, Responsibility and Rights: Feminist Ethics and Social Theory*, eds. Hilde Nelson and Robin Fiore (Lanham: Rowman and Littlefield, 2003), 90.

> standards of diachronic rationality, which is a cross-temporal assessment of rationality.[26]

This case does indeed typify ones that responsibilists have raised.[27] Dougherty even sharpens this challenge by imagining that he suggests to Craig "that he read a couple books which I say show that the arguments of young-Earth creationists are seriously flawed," and that "he refuses to do so." [28] If we allow the dynamic aspect of evaluating Craig over some interval of time, this refusal to inquire into the matter further by consulting readily available books and arguments concerning geologic time, etc. seems intellectually irresponsible and something an intellectually virtuous agent would not do. If we further trust the intuition that there is a change (a drop) in our evaluation of either or both of Craig himself, *qua* epistemic agent, or of the warrant he has for his belief, then the case appears to undermine the Identity Thesis. For again, "evidentialism requires the change in *epistemic* status of belief issue from a change in *evidential* status,"[29] and here it seems Craig has no substantial new evidence, having ignored rather than investigated the challenges raised by his friend's testimony.

But Dougherty, having now nicely identified the challenge that the Craig Case presents to evidentialism, goes on to give evidentialist explanations of it, explanations that proceed without a need for the 'diachronic picture,' and that apply a kind of error theory to those attributions of intellectual irresponsibility which on first glance had looked formidable. Initially we were supposing Craig's belief to be justified by the evidentialist standard, even if only because the homogenous culture he was raised and educated confirmed this belief at every step and he never sought out any potentially disconfirming evidence. But on the first of these three alternative interpretations we are mistaken to think that there is no change to Craig's evidential status. There *is* a loss or drop: After Craig's encounter his belief *no longer* fits the evidence as it previously did. Recognition of serious peer disagreement eventuates in the loss of whatever positive epistemic status of Craig's belief that evidential fit is supposed to supply, but explains this by a *corresponding* drop in his evidential status or 'fit.' "Evidence of evidence is evidence," so that "at the time at which Craig became aware that he had testimony from a known reliable source that some of his beliefs were false, at that moment he had *evidence* that his beliefs *were* false."[30] Thus, this loss of positive epistemic status

[26] Dougherty, "Reducing Responsibility," 5.
[27] For others see especially Baehr, "Evidentialism, Vice, and Virtue."
[28] Dougherty, "Reducing Responsibility," 6.
[29] Dougherty, "Reducing Responsibility," 5, emphasis original.
[30] Dougherty, "Reducing Responsibility," 7.

of his belief "is perfectly explicable in terms of synchronic justification"[31] because Craig is interpreted as having gone *from* synchronically justified ('epistemically rational') to synchronically unjustified ('not epistemically rational') in holding his belief. Hence by this first explanation, "If his beliefs don't change at all we have something to explain our inclination to condemn his belief in the lack of synchronic justification."[32]

This first explanation functions so as not to allow responsibilists to run cases in the way they want in order to directly challenge the Identity Thesis, and Dougherty believes that in many cases this explanation may offer the most natural interpretation. He is nevertheless willing to concede that responsibilists *would* be able to describe the case so that it is clearly assumable, or is stipulated, that Craig remains synchronically justified in accepting the creationist proposition *after* the encounter with his friend. Thus Dougherty allows his first response to be treated as something of an aside, acknowledging that dialectically the greater weight of argument over issues between the evidentialists and the responsibilists falls upon the success or failure of the latter two of his three evidentialist explanations of the Craig Case.

These second and third explanations take a markedly different tact than the first. According to these, Craig maintains the 'fit' between his level of belief and his total evidence even post-encounter, and so long as this is true of him then he is doing perfectly well *qua* epistemic agent.[33] Craig is indeed culpable in *some sense* for failing to pursue inquiry and/or discounting the testimony of the epistemic peer who disagrees with him, but the second explanation is that what he displays is only an 'instrumental' or 'prudential' irrationality. "It is instrumentally irrational for Craig not to read the books, since if he wants to get the truth, and we supposed he did, he ought to read the books, since he has evidence that reading the books will get him to the truth. The cost is low, the potential payoff, we may assume, is high."[34] The second explanation leans heavily upon the point that there is no general philosophical principle as to what determines whether acquiring some true belief is worth our effort attention; expending effort on inquiry and taking an attitude towards a proposition always involves personal costs and trade-offs, and since there are many truths not worth knowing, any answer to what determines whether some truth is worth our attention engages personal/practical interests and/or moral concerns. Although he doesn't use these terms, Dougherty might be

[31] Dougherty, "Reducing Responsibility," 7.
[32] Dougherty, "Reducing Responsibility," 10.
[33] Earl Conee and Richard Feldman, ed., *Evidentialism: Essays in Epistemology*.
[34] Dougherty, "Reducing Responsibility," 10.

thought to be accusing the responsibilists of conflating theoretical and practical reason, or again of presupposing a strong voluntarism about belief in ascribing 'intellectual' responsibility and irresponsibility to agents.

But what if it is responded that this explanation isn't plausible because truth is, after all, not some merely pragmatic goal but a prime (perhaps even intrinsic) epistemic good, and it looks as though Craig's faulty way of maintaining his challenged belief shows that he lacks appropriate motivation to believe truly? Then there is a third explanation which an evidentialist might invoke: "But this also is either a matter of instrumental irrationality if he has a sufficiently strong desire to believe truly on this matter or a purely moral one if he does not have such a desire."[35] There is a kind of irresponsibility operating, but it is purely moral and not intellectual irresponsibility. This relates back to the earlier point about 'culpable ignorance' when there are costs to information. Dougherty doesn't take himself to be arguing that there is no such thing as culpable ignorance in matters of belief, but rather that "being in a state of ignorance is, when irresponsible, morally irresponsible or instrumentally irrational."[36]

The evidentialist reductionist does not need to argue that these three explanations are mutually exclusive, but they do need to show that they are all that are needed in the Craig Case, and that it's unlikely that we will find other real-world cases that cannot also be better explained in at least one of the ways indicated. Dougherty holds that simplicity argues for reducing normative categories wherever possible, and that if epistemic responsibility *can* be adequately explained through any one of the three proposed strategies then we needn't go looking for a different kind of explanation.

4. Interpreting the Craig Case

Let's now turn to critically evaluate each of Dougherty's proposed explanations of the Craig Case in turn. His first explanatory strategy *allows* that there is a loss[37] in the epistemic status of Craig's creationist belief after the encounter with his friend and epistemic peer, but it explains this loss in terms of a *corresponding* loss of the evidential fit that his belief enjoyed prior to the encounter. Hence it explains the drop in status wholly in terms of loss of synchronic rationality or evidential fit.

[35] Dougherty, "Reducing Responsibility," 10.

[36] Dougherty, "Reducing Responsibility," 4.

[37] Dougherty appears to differ from Feldman in allowing that there are degrees of belief, but it seems to me that he applies this inconsistently to his epistemology of disagreement if he still assumes as Feldman does that a person's justification for a belief is fully defeated by the awareness of disagreement, or that no belief is defeated by this awareness.

I would concede that there are in principle certain cases for which each of the three explanatory strategies might work. But in what kinds of cases, specifically, and if in the Craig Case, then why? Dougherty makes no effort to clearly identify what the moral fault or the prudential aim would be construed to be in the Craig Case, for instance. With the first explanatory strategy, it is clearly the evidentialist's burden to provide not a mere sketch or suggestion but a full accounting, a *theory of evidence* illuminating the actual basis for claiming that Craig's belief was synchronically rational *up to* this one encounter, but synchronically irrational thereafter. Otherwise invoking this explanation appears *ad hoc*: What is the independent motivation for viewing *this* one particular piece of testimony or *this* one encounter with an epistemic peer as being what 'tips the scale' or 'crosses the threshold' between synchronic rationality and irrationality?

Dougherty and I would agree, I think, in taking recognition of serious peer disagreement as having epistemic consequences for agents. Dougherty does not seem entirely comfortable with those currently-popular epistemologies of disagreement that allow only the three doxastic attitudes of full belief, disbelief, or suspension, and that consequently tend to leave only the options of a 'no defeater' or a 'full defeater' view of the epistemological significance of serious peer disagreement. I would concur with rejecting these assumptions, if that is indeed what Dougherty intends. But allowing that epistemic-peer disagreement often results in a *partial* defeater,[38] while likely making the description of the focus-case more realistic, would correspondingly also make it more difficult to support the assertion that Craig passes over this threshold from being synchronically justified to being synchronically unjustified in holding his creationist belief, at just the point at which he receives his peer's contesting testimony.[39] As a further aside, epistemic 'peerhood' even seems impossible to define apart from admirable shared habits of inquiry and a normal level of intellectual motivation to achieve a range of socially shared and acknowledged epistemic goods. Thus I think that attributions of evidential fit with 'total evidence' are implicitly *parasitic* on attributions of adequate investigatory habits, suggesting again that bare synchronic rationality maximization is not a fundamental criterion of epistemic value. Note that responsibilists also often say much the same thing about bare or 'thin' attributions

[38] As does Michael Thune, "'Partial Defeaters' and the Epistemology of Disagreement," *The Philosophical Quarterly* 60, 239 (2010): 355-372.

[39] Under the influence of evidentialist assumptions, "the debate about the epistemological significance of disagreement is largely focused on whether a belief's justification is retained or lost wholesale," yet a more adequate "account of the epistemological significance of disagreement should leave plenty of room for cases which result in the partial defeat of beliefs." (Thune, "'Partial Defeaters,'" 356 & 372.)

of reliable genetically-endowed processes (faculty virtues) made by austere versions of reliabilist externalism that neglect the importance of personal justification, both synchronic and diachronic.[40]

This suggests a further line of argument. The first explanatory strategy illustrates for us how evidentialists want to treat Craig's belief on the model of acceptance of an atomistic proposition, and want to treat the total evidence he has for that proposition in the way we would treat the evidence for any everyday, flatly empirical assertion. Yet in most matters of any importance in relation to a person's moral, political, philosophical, or religious beliefs the evidence is diverse, subtle and complex; what the agent finds salient in this diverse evidence-base, and therefore what constitutes the total evidence he 'has' that bear upon it, will vary over time. The reasoning that supports propositional acceptance based upon such evidence is *cumulative in form:* It is not just a matter 'total' evidence, but of *totalizing* or on-balance reasoning. *Cumulative case arguments* are ones that agents approach with cognitive *strategies* that vary greatly in their sophistication and in their appropriateness to the subject matter before them. For example, a figure like Craig (who hasn't examined what we would call 'hard' empirical evidence) likely relies pretty heavily upon Paley-esque arguments by analogy and arguments by appeal to Biblical authority. It is these I imagine he would cite if pressed to offer defense of his special creationist belief. But both of these inductive strategies are notoriously difficult to provide objective/probative measures for. Logic tells us that with analogical argument and appeal to authority, we need to carefully distinguish our evaluation of the argument's inductive strength from our evaluation of the truth of its premises; yet many or most people are not logically proficient enough to notice let alone to carefully apply this distinction between inductive *strength* and *cogency* in their reflective weightings of these kinds of arguments.

The upshot here is that an evidentialist attributor of the first explanation must first be *in a position to know or reasonably judge* that Craig's encounter with the friend and epistemic peer really does tip the scale from synchronically justified to unjustified – from *on-balance* fitting to *on-balance* not fitting his total evidence; but the cumulative or all-things-considered nature of the agent's inference in such cases strongly suggests that the evidentialists are here setting themselves a task that cannot be met. The problem of operationalizing 'evidential fit' to a degree that can evade the *ad hoc* objection will be still more difficult whenever broad-scale hypotheses or theories confront one another; now the choice between them is generally characterizable as *inference to the best explanation*. Cases of *inference to*

[40] See George Streeter, "Virtues of Inquiry and the Limits of Reliabilism," *Social Epistemology* 20, 1 (2006): 117–128.

the best explanation are again common in science, philosophy and religious metaphysics, but they place a further special burden on the evidentialist who wants to apply Dougherty's first explanatory strategy to the Craig Case: *produce the further weightings* that allow us to support the claim that an agent like Craig was really on-balance synchronically justified in his belief at time T1 but not so at T2, post-encounter with a proponent of a contrary large-scale hypothesis or theory.

Another major worry about the satisfactoriness of the first evidentialist explanatory strategy also arises from Dougherty's rather artificial presentation of Craig as an agent previously situated in near complete cultural isolation from the scientific tradition of our day. The author describes this as a 'harmless idealization'[41] and says his account could easily be adapted to more realistic ones and still run this first explanation. I doubt that this is true, however. For Craig's epistemic situation to be one that genuinely invites interpretation along the lines of the first explanation, such an idealization seem needed.[42] In a case where the individual has already been aware of scientific ideas which challenge his special creationism, it will be all the more difficult to motivate the claim that this one *further* peer-encounter or one *additional* piece of testimonial evidence results in that agent's belief sliding from properly to improperly fitting his total body of evidence.[43]

The second and third explanations that Dougherty offers, by allowing us interpretations on which Craig continuously fulfills the evidentialist's condition of being synchronically rational, are better keyed to the issues at stake between evidentialist supporters of Identity and their critics. The responsibilists now appear in a better position to argue that the deserved criticisms of Craig *qua* epistemic agent aren't captured on assumption that synchronic epistemic rationality is the sole source of legitimately epistemic norms, but rather by something like a failure in due care and diligence on Craig's part. But in reply to this charge, Dougherty is arguing that the evidentialist might succeed in saving the Identity Thesis by

[41] Dougherty, "Reducing Responsibility," footnote 7.

[42] The unrealisticness of this idealization is the one small exception I mentioned earlier to allowing the Craig Case as a shared focus for debate.

[43] Note that in the page 5 quote, above, Dougherty writes that "I'd often recommended books challenging his views" in the past. In this more realistic glossing of the case, I'd hold, the first explanation is *clearly implausible* because the newest encounter is unlikely to dramatically result in moving the agent's belief from on-balance fitting to on-balance not fitting total evidence. The encounter with a disagreeing peer is then clearly an important piece of evidence (and serious peer disagreement, I agree with Dougherty, deserves to be taken seriously). But I am skeptical of an ideal attributor providing the single objective weighting of these diverse evidences needed to say that an agent was synchronically rational at time T1, but synchronically rational at time T2, post-encounter.

sticking to their guns and saying, first, that in remaining synchronically justified Craig is 'doing fine' as an epistemic agent, and second, that what the responsibilists describe as Craig's intellectual irresponsibility will always on closer analysis be seen to be an extra-epistemological evaluation. More formally, the suggestion that the responsibilists are confusing intellectual responsibility with moral blameworthiness is first broached by Dougherty, as we noted earlier, through the question of what determines whether a particular truth is worth one's attention. From there, his line of reasoning leading to the second and third explanations is formalized as a constructive dilemma:

> Either there is some interest at stake in knowing or there is not. If there is not, then there is no irresponsibility. If there is, it is either the inquirer's interest – or someone else's interests are at stake – in which case it is a moral shortcoming.[44]

A constructive dilemma is a valid argument form.[45] This dilemma I'll argue is unsound, and I'll approach it by 'grasping' the second of its two horns.[46] The dilemma as just quoted requires further clarification, however, since its second horn is itself intended to be disjunctive, leading directly to Dougherty's second and third explanations:

> 2nd Horn: if it's the inquirer's interests at stake then the irresponsibility or blameworthiness is really just a matter of instrumental irrationality, and if it's someone else's interests at stake then it's really a matter of moral irresponsibility.

First, some general comments regarding assumptions underlying Dougherty's constructive dilemma. One mistake that I think is made here is to cast these issues in such an individualistic philosophical idiom. Knowledge and knowledge

[44] Dougherty, "Reducing Responsibility," 1.

[45] A constructive dilemma is a valid argument of the form (if P then Q) and (if R then S); P or R; therefore, Q or S. There is actually an embedded second dilemma attached to the second horn: "If it concerns the interests of others, then either I have a duty to promote their interests or I don't. If I don't, then I'm doing nothing irresponsible in not being scrupulous. If I do, then the irresponsibility is clearly moral." So if it concerns the interests of others, then either I'm doing nothing irresponsible, or my irresponsibility is clearly moral. My response, which consists of grasping the horns of the constructive dilemma, encompasses a response to this second, embedded dilemma as well.

[46] The first horn – that where there is no interest at stake in a particular proposition being believed, there can be no question of irresponsibility, intellectual or ethical – I won't challenge. But there is a worry about providing a mantle for vices of inattention, which I do not consider to be all and only moral vices. The issue is properly that one can never be intellectually blameworthy for not believing or not disbelieving a proposition regarding which they have no evidence that a morally and intellectually virtuous agent would be presumed to have; it is not that lack of attention entirely insulates an agent from epistemic appraisal.

attributions serve social functions, such that agents are members of communities of inquirers, and knowledge-attributions 'genealogically'-considered serve social functions of marking reliable testifiers. Another problem that needs to be avoided in the framing of the dilemma is treating assumptions stemming from the fact/value dichotomy of Dougherty's Model [1] as if they were analytically true (true by definition). This cannot be done without begging some of the most interesting questions at issue in the present debate. I will highlight what I see as mistakes of this kind as they crop up with the nothing but 'instrumental irrationality' and the nothing but 'moral irresponsibility' reductions.

Let's start with the third, or 'moral irresponsibility' strategy. Dougherty's point that we sometimes confuse moral with epistemic appraisal is again a valid one. But our burden is only to say why in the Craig Case and similar cases this third explanation does not look very plausible or philosophically satisfying. Firstly, the responsibilists think that we should consider Craig *intellectually irresponsible* for dismissing his peer's disagreeing testimony and suggestion to re-open inquiry into the grounds for his belief, but not necessarily or obviously *morally irresponsible* for doing so. The evidentialist who appropriates the third explanation will be saying asserting exactly the opposite; I doubt that evidentialists have consensus intuitions in this matter on their side. This is a minor point in itself – what the evidentialist is offering is, after all, an *error theory*, and *could* encompass the error of even what seem to me strong common-sense intuitions to the contrary. What it does is only to raise the bar on the need to show the clear superiority of their way of explaining the case. Secondly, it seems that far too much of the weight of argument with the third explanation flows *merely* from identification of the 'interests at stake' in a particular instance of knowing with interests of 'others' besides the agent him or herself. That every self-regarding consideration is a non-moral consideration, and every other-regarding consideration is a moral one, is not an assumption that I think many ethicists share, whatever normative ethical account they hold. Secondly, far too much of the weight of argument with the third explanation flows merely from association of the dynamic aspects of belief acquisition and maintenance with overt 'behavior related to the formation of belief.'[47] This term 'behavior' threatens to carry practical or moral assessment by definition, potentially placing it all within the realm of practical reason with one fell swoop. But of course much of what the defenders of character epistemology and a diachronic picture of epistemic evaluation are concerned with are intellectual dispositions, habits and problem-solving strategies, of which overt behaviors are outward manifestations. Conduct and character are strictly correlative, as Dewey puts it; virtue theorists

[47] Dougherty, "Reducing Responsibility," 3, Feldman, "The Ethics of Belief," 90.

would anyway be among the *least* likely fellows to focus on conduct to the exclusion of character. So this term 'behavior,' while convenient for the evidentialist to employ, misses its mark by a good distance in describing the responsibilist approach.

Thirdly, if I claim that an agent lacks the motivational set needed to distinguish aesthetic objects from mechanical ones, my judgment of the agent is not *itself* an aesthetic judgment but a judgment of agential competence *with respect* to this target domain. Relatedly, an evidentialist characterization of Craig's blameworthiness as relating to his moral rather than intellectual agency seems to confuse two things. It may well be moral considerations, such as the impact that an agent's highly ethnocentric belief might have on others when he acts upon it, that initially draws philosophic or social scientific researchers to an interest in the case, and to scrutinize habits of inquiry that the agent manifested in acquiring or maintaining this belief. But once so attending, the focus easily shifts to epistemic evaluation simpliciter, where the standards utilized are intellectual standards involving consideration of whether the agent manifested normal intellectual motivations, utilized effective or ineffective cognitive strategies for problem-solving, avoided known cognitive biases and fallacious tendencies in reasoning, etc. This latter evaluation will be clearly epistemic to the extent that what the presence (or absence) of normal desire for true belief and strategic efforts at inquiry is *salient in explaining* isn't the agent's blameworthiness in *acting* upon the belief, but simply why the agent in this instance was or wasn't successful in achieving distinctively *epistemic aims* such as true belief, etc. This is then a distinct evaluation of the quality of the agent's inquiry and of the competence and performance of the agent *qua* inquirer. If this weren't the case, then presumably the proponent of the third explanatory strategy would have to hold that all attributions of 'epistemic credit' on credit accounts of knowing are disguised moral assessments as well.

Enough has been said regarding the claim that if the interests at stake in knowing in a particular case are not the agent's own, then evaluation of that agent's habits of inquiry is really only disguised moral evaluation. But let's look at the other conjunct of the second horn of Dougherty's dilemma, the one which states, "If it pertains to one's own interests, then the irresponsibility at hand is easily explained in terms of practical irrationality."[48]

In my view the Craig Case and cases like it will not plausibly be adequately explained on the other two explanatory strategies, and so the ability of this 'practical' or 'instrumental irrationality' explanation to catch all remaining is crucial to the support of the Identity Thesis. Explaining the agent's failure of due care and

[48] Dougherty, "Reducing Responsibility," 7.

diligence with counter-evidence will be necessary if we are to be convinced that the Identity Thesis is true and that there aren't real-world cases of apparent intellectual irresponsibility that cannot be explained in one of these three ways.[49] Yet Dougherty's reasons for claiming that diachronic assessments of the intellectual irresponsibility of the agent amount only to assessments of 'practical' or 'means/end' or 'instrumental' rationality/irrationality, strike me as obscure and dubious. Yes, our intellectual lives are always caught up in the balancing of the goals of accumulating interesting true beliefs and avoiding false ones; this involves both valuation and risk, a point that pragmatists like Haack have often highlighted. There are also, to be sure, issues concerning tradeoffs between the cost of information in terms of energy and attention, and the constant need for active fair-mindedness. I agree with Dougherty that his points about how practical interests and values may inform issues of attention, etc. serve as a corrective to any veritists who would claim "a general, impersonal duty to seek truth as such."[50] But these points I see going nowhere in regards to showing that all choices of attention and all investigative strategies in pursuit of true belief, knowledge, understanding, etc. are nothing but moral directives or else matters of prudential/instrumental rationality. If doing fine as an epistemic agent were really the passive affair the synchronic account seems to suggest, there would be no response to the 'updating problem' with inductive knowledge, which in turn would be a steep concession to scepticism. There may be no general answer to the question of how long you remain justified in your belief that the car you parked on the street yesterday, or the child you parked in the playroom before the game started, is there yet. But the reasonableness of occasionally getting off the coach and checking ('updating one's evidence') is not merely pragmatic/moral – it's that, to be sure, but one's level of active searching for updating information is clearly *also* a matter of what rational confidence you can have for those inductive beliefs, and what good reasons you could offer if challenged in them. Evidential fit is presumably supposed to supply reflectively 'good reasons' that one can discursively offer as grounds for the belief. But if the agent hasn't been active in updating information when needed in order to maintain rational confidence, or put any effort into inquiring into counter-evidence to their belief once presented with it, then that agent's reasons aren't going to wash when someone asks them why they (still) believe it, are they?

---

[49] The language of intellectual vices seems well-adapted to providing detailed characterizations of the cognitive heuristics as cognitive studies are revealing to us, without treating agents who fall afoul of the standard of evidential fit as simply 'dysfunctional' agents, as Dougherty appears to want to treat them.

[50] Dougherty, "Reducing Responsibility,"10.

It is noteworthy that Dougherty uses 'prudential,' 'instrumental' and 'practical' synonymously, in order to mark a sharp separation from matters epistemic. But why can't it be both instrumentally and epistemologically evaluable, where true *belief* is valued intrinsically and true *beliefs* serve many a practical goal? There may well always be the possibility of additive personal value from having true beliefs that are relevant to action, or to satisfaction of personal desires, but this doesn't show that their personal value *excludes* their status as an epistemic good. A decision to refrain from some activity "does not constitute a rejection or a denial of the norms that govern it"[51]; neither does choosing to engage in that activity render its assessment wholly pragmatic.[52] An agent who *while* engaged in inquiry forms reflective beliefs by way of methods or strategies that aren't reliably truth-conducive seems to be engaging poorly in epistemic activities. Efforts to be actively fair-minded are obviously intimately involved in the improvement of agent reliability in areas of contested belief. It seems that if the efforts to be actively fair-minded towards counter-evidence to one's belief is 'easily explained in terms of practical irrationality' it is only because instrumental or means-end redescriptions are *always easily available* when aim-pursuit is involved. Its availability in and of itself does little to show that such instrumentalist descriptions are *adequate* to their subject-matter, however. Their availability as descriptions doesn't mean they capture all of the relevant aspects of rationality involved. Cases where the means are partly *constitutive* of the end, as seems to be the case with intellectual as with moral virtues, are among the many cases in which instrumentalist redescription provides too narrow a conception of the rationality-connected concerns. Nickerson identifies at least eight distinct conceptions of rationality, which he sees as 'aspects' dependent upon explanatory interests. These are rationality as:

- consistency with self-interest
- pragmatic adaptiveness
- consistency of actions with preferences or goals
- optimal analytic choice behavior

[51] Michael-John Turp, "Naturalized Epistemology and the Normative," *Forum Philosophicum* 13 (2008): 343-355.

[52] "We do, after all, decline to investigate the truth of a great number of propositions without this implying that we thereby violate in the norms of judgement. We might do so because the propositions are trivial, or because the investigation would be inappropriately time-consuming, or for moral, prudential, political or aesthetic reasons …. However, a decision to refrain from some activity does not constitute a rejection or a denial of the norms that govern it. If one decided that [some project] were a reasonable project, then quite plausibly one *would* be governed by the applicable epistemic norms, as in all one's epistemic endeavors." (Turp, "Naturalized Epistemology," 352.)

- satisfycing
- conformity to norms
- reflectiveness
- responsiveness to reasons.

Those most clearly epistemological among the eight aspects of rationality would be rationality as reflectiveness, and as responsiveness to reasons. The responsibilist's interests most clearly fall under rationality as reflectiveness – "a matter of attitude and intent – love of truth, willingness to examine issues from various points of view, active fair-mindedness"[53]; but the responsibilist makes no pretension that this is the *whole* of rationality as related to the achievement of epistemic goods. The evidentialist's interests most clearly fall under rationality as responsiveness to reasons if we restrict 'responsiveness' to the passive and synchronic, as they would. They *do* claim that this is the whole of epistemic rationality, and that rationality as reflectiveness, because it often involves active and not merely passive fair-mindedness, reduces to rationality as consistency of actions with preferences or goals. This seems unmotivated. How much effort rationality requires that an agent expend is an open question; clearly the rational requirement to seek evidence must be tempered by the cost of obtaining information and by the recognition of practical limits on what people can be expected to do. But it is far from clear that 'epistemic rationality' consists exclusively in being responsive or fair with whatever evidence one *happens* to encounter. "A more active interpretation would have rationality require that one put significant effort into seeking evidence … active search for evidence, especially counterindicative evidence, is a key aspect of some conceptions of what it means to reason well."[54]

Recalling a passage noted earlier, "maybe the problem is that we think Craig doesn't care enough about the truth. But this also is either a matter of instrumental irrationality if he has a sufficiently strong desire to believe truly on this matter or a purely moral one if he does not have such a desire."[55] This strikes me as another false dilemma. Briefly, if the agent's performance in gathering and weighing evidence is thought to be irresponsible for want of skills rather than sound intellectual motivation, then the sense of 'instrumental' irrationality that could be applied would clearly be consonant with epistemic evaluations rather than representing merely practical or pragmatic irrationality.[56] Yet the case only appears to get worse

[53] Nickerson, *Aspects of Rationality*, 25.

[54] Nickerson, Aspects of Rationality, 141.

[55] Dougherty, "Reducing Responsibility," 10.

[56] To illustrate further, there are epistemologists who conceive first-personal epistemic rationality as a *kind* of instrumental rationality in the service of one's distinctively cognitive or epistemic

for this explanatory strategy if we hold true belief to be an aim of intrinsic value. Nor need we agree, as the second horn of the newest dilemma supposes, that if Craig's intellectual motivations *are* what is in questions – that is to say, if indeed we find that lack of a normal level of intellectual motivation leads to his unreliability either generally or in a particular case, that this is necessarily to make a moral assessment of him. The previous objections to the second explanatory strategy I think already work to dispel this contention.

In summary, if the *ad hocness* worry was strong with the employment of Dougherty's first explanatory strategy, that worry is greatly magnified with this third strategy, which purports to show that evaluations of an agent that *aren't* moral evaluations but that do go beyond 'evidential fit', can always be treated as some 'other sort of normative failing'. 'Instrumentally' and 'prudentially' irrational are the catch-all terms, but it appears that pursuit of perfectly respectable epistemic ends are simply being *re-described* in this manner in order to save the reductionist theory. Yet the real point of contention may not be these issues about differences between epistemic and prudential assessments, but about sharp differences in what Kvanvig[57] calls the 'idealizations' embodied in different epistemological theories. In the way that virtue responsibilists idealize personal justification, it depends not just on the information the agent in fact recognizes, or has available to her, but also information that person *would* have obtained given sound epistemic motivation and a baseline degree of cognitive competence on that person's part. The standard of competent due care and diligence that an agent's performance is judged against is not restricted to a passive conception of fair-mindedness, but it does not need to

goals; Thomas Kelly dubs this the *instrumentalist conception* of epistemic rationality, and it is exemplified in a framework of naturalized epistemology by thinkers like Hilary Kornblith. This view has some plausibility, but assessing Craig's shortcomings as a kind of instrumental irrationality using this model would confirm rather than deny its standing as an epistemological assessment of the agent. If this *isn't* what Dougherty intends, then I suspect he is just smuggling in at this point the quite radical Conee/Feldman stance "if there is an aim for belief, or a norm for belief, it is evidence, not truth … if a person has strong evidence for a false proposition F she should believe that falsehood" (Conee, Feldman, *Evidentialism: Essays in Epistemology*, 184), and that "a person who irrationally believes a lot of truths is not doing well epistemically. In contrast a person who forms a lot of false beliefs rationally is doing well epistemically." (184) But, as Pascal Engel has noted, to purport to supplant knowledge, and understanding as the telos of the life of the mind with constant, perfect synchronic rationality is a radical doctrine that finds no motivation aside from a host of assumptions not shared by evidentialism's critics (Pascal Engel, "Review of *The Architecture of Reason*, by Robert Audi / *Evidentialism*, by Earl Conee and Richard Feldman," *Disputatio* 20, 1 (2006): 349-358, http://disputatio.com/articles/020-5.pdf, accessed October 2010).

[57] Kvanvig, "Propositionalism and the Metaphysics of Experience."

insist upon unrealistically large efforts at inquiry either. That level is determined by the context of inquiry itself, including the nature of the claim, and the kinds of investigative and reasoning strategies that bear upon it.[58]

5. Closing Remarks

Trent Dougherty's paper highlights important issues that deserve considerably more attention. Issues about active fair-mindedness and the cost of information is one such question responsibilists would like to see discussion expanded upon. This paper has replied directly to his dilemma for responsibilism, and to his key reductionist claim that one or another of his three proposed explanatory strategies will always better-explain cases like the Craig Case than will any diachronic or dynamic concept of epistemic irresponsibility. I conclude that none of the three explanations Dougherty offers in the Craig Case is as plausible as he thinks they are, and that neither individually nor collectively do they provide evaluations *as* plausible or informative as one invoking person-level motivation and/or habits of inquiry. It isn't an error or illusion: Craig really *isn't* doing as well *qua* epistemic agent as Dougherty and other evidentialists would have us suppose. Moreover, the merits of a reductionist urge in epistemology, as Dougherty presents it, is considerably over-wrought. Reductionist philosophical approaches that claim motivation from a desire for theoretical unification often not only fail at their task, but actually preclude the possible realization of greater theoretical unification. Evidentialists like Feldman claim to want to develop as purely internalist a conception of epistemic justification as possible. The evidentialist insistence on reducing epistemic normativity to a purely synchronic standard, and the perpetuation of the longstanding debate over internalism and externalist about knowledge, I submit, is a case in point.

The objections we have made here to Foley's and Dougherty's Identity Thesis and its confusions of *evidential* status with *epistemic* status largely agree with those that Richard Foley has previously made; this makes relevant the lesson Foley draws from his own discussions of internalism, the lesson of the need to properly separate the theory of knowledge from the internalist account of justification. The key

[58] Dougherty could argue directly against such agent-focused idealizations, but the plausibility of the explanation in terms of Craig having exhibited instrumental/practical *instead* of intellectual/epistemic blameworthiness cannot lean on the assumption that propositionalism is true without begging all of the interesting questions at stake here. Propositional and doxastic justification are indeed difficult to reconcile, whatever epistemological orientation one adopts. But virtue epistemologists do not think the evidentialists are fair in imposing the self-serving propositionalist view over the debate. Etiology matters in epistemology, and diachronic factors are substantially a matter of a belief's causal etiology.

lesson for Foley is "the corrupting consequences of the assumption that there is a conceptual tie between epistemic justification and knowledge."[59] This "Unfortunate Assumption," as he calls it – the assumption that the conditions that make a belief justified are by definition conditions that turn a true belief into a good candidate for knowledge – is "needlessly limiting": "It discourages the idea that there are different, equally legitimate projects for epistemologists to pursue" and it inevitably distorts *both* the project of trying to provide and analysis of knowledge, *and* the project of understand epistemic responsibility or rationality its importance to us. Now, what Dougherty calls the Identity Thesis is a product of the Unfortunate Assumption and appears completely unmotivated apart from it. What commitment to that assumption produces in the present instance is an artificial separation between the standard of evidential fit or synchronic rationality and our everyday assessments of each other's opinions, which as Foley points out "tend to emphasize whether we have been responsible in forming our beliefs rather than whether we have satisfied the prerequisites for knowledge."

'The remedy' which Foley prescribes, and which I here suggest we prescribe for defenders of the Identity Thesis, is that they not resist the philosophical reasons pushing us to a 'separation' of the two projects. The prescribed remedy is for them to jettison the Unfortunate Assumption and all associated attempts to forge a necessary link between evidential justification, and knowledge possession, and then to develop an account of synchronic rationality as one distinctive sense of personal justification or responsibility, and one among several epistemic factors or 'springs' of epistemic value.

John Greco's recent book, *Achieving Knowledge: A Virtue-Theoretic Account of Epistemic Normativity*,[60] provides quite more direct arguments against the evidentialist understanding of epistemic justification and epistemic normativity than I have tried to give here. Knowledge-relevant normative status is not exhausted by the facts about one's evidence. The view that it is exhausted by evidential fit, he argues, suffers from a "psychological plausibility problem" and is undermined by contemporary cognitive science, and "the prospects for evidentialism about knowledge look bleak."[61]

But in order not to conclude with this criticism of reductionism in epistemology, let me simply propose a model of diachronic evaluation of the competence and

[59] Richard Foley, "Justified Belief as Responsible Belief," in *Contemporary Debates in Epistemology*, eds. Matthias Steup, Ernest Sosa (Malden, Oxford, Carlton: Blackwell, 2008), 313-326. All quotes from pages 314-315.

[60] John Greco, *Achieving Knowledge: A Virtue-Theoretic Account of Epistemic Normativity* (New York: Cambridge University Press, 2010).

[61] Greco, *Achieving Knowledge*, 68.

performance of inquirers in the pursuit of epistemic aims or goods. The issue between Dougherty and the core responsibilists is not about what has been called 'pragmatic encroachment' upon epistemic norms; but if like Feldman one tries to construct as an account of epistemic justification as internalist as possible, the issue at hand might well be seen as a challenge of 'diachronic encroachment.' On the following chart reflects key concerns with diachronic evaluation that I think are today shared among social, feminist, and character epistemologists, and that constitute responsibilist research programs in the sense of programs that put epistemic responsibility into a central place in the theory of knowledge. "Normative epistemological assessment need not be restricted to judgments of whether particular claims or beliefs are good or bad, epistemically speaking. Normative epistemological assessment can also take the form of developing models of how to *practice* good inquiry … action and inquiry are central to the concept of an epistemically responsible agent." [62] The diachronic are evaluable in at least three basic ways: morally, pragmatically, and intellectually. We begin with our interests in diachronic evaluations, and then examine them 'in light of' moral, pragmatic, or intellectual and epistemic concerns. We do not precategorize synchronic and diachronic norms as wholly one or another, but allow the agent's context and our own interests as evaluators to focus our assessment in one or more directions. Diachronic evaluations are sometimes backwards and other times forward-looking, and this is true of each of the three ways just mentioned. So this model yields six basic types of evaluation, to which on the chart I have added basic 'prompting questions,' to give you a feel for how I'm thinking about each. In some cases I have further separated out prompting questions that motivate social epistemological evaluations from those that motivate individual-agent epistemological evaluations. I use my usual terminology in this, referring to these as *zetetic* evaluations, because the diachronic as here understood identifies inquiry-directed motives, habits, and activities, or *zetetic* considerations.[63]

[62] Grasswick, "The Impurities of Epistemic Responsibility," 91-92.

[63] Thanks especially to Trent Dougherty, Phil Olson, Heidi Grasswick, Nancy Daukas, Christian Miller and Timothy Chappell for stimulating discussion, as well as to many other discussants at *JanusBlog: The Virtue Theory Discussion Forum*, http://janusblog.squarespace.com.

Responsibility Unpacked: Doxastic Norms & the Logic of Diachronic Evaluations

Epistemological	**Practical**	**Moral**
Individual: The etiology of belief as relevant to doxastic justification of beliefs, and hence to epistemic credit attributions and to final value (the value associated with achievements of all kinds). Individual *zetetic* prompting question: "What did I do that my belief was grounded on solid and adequate evidence?" Social: Genealogical narratives illuminating the *axiological* (normative/value or authority-conferring) and social functions of epistemic concepts and attributions within our overall cognitive ecology. Social *zetetic* prompting question: "What functions, descriptive and/or normative, does this concept serve within epistemic practices and communities, and how does its history illuminate these functions?"	Individual *zetetic* prompting question #1: "What did I do to ensure that the beliefs and attitudes upon which my actions were based were formed with a degree of responsibility fitting the gravity of the decision?" Individual *zetetic* prompting question #2: "Did I exhibit the kinds of non-epistemic deliberative virtues – e.g. friendliness or sincerity that conduce to a healthy deliberative environment and good epistemic practice?" Social *zetetic* prompting question #1: "Was a healthy deliberative climate present during the time of my inquiries and deliberations?" Social *zetetic* prompting question #2: – "Were the background conditions during the time of my inquiry and deliberation such that good epistemic practices could be pursued?"	Individual *zetetic* prompting question #1: "What did I do to ensure that the beliefs and attitudes upon which my actions were based were formed with a degree of responsibility fitting the gravity of the decision?" Individual *zetetic* prompting question #2: "Were my personal interests and motivations consistent with ethical virtue and with a good life?" Social *zetetic* prompting question #1: "Was there epistemic justice or injustice displayed in the group institution's deliberative practice in question?" Social *zetetic* prompting question #2: "Was the division of epistemic labor fair, and were varied perspectives fairly included in this group or institution's deliberative practice?"
Backward-looking **Forward-looking** **v** **Epistemological**	**Practical**	**Moral**
Improving one's or one's group's epistemic situation, intellectual motivations, and/or problem-solving strategies. Individual *zetetic* prompting question: "Is my epistemic situation good enough for me to understand the matter, or to make a sound judgment in accepting or rejection a proposition; or should I first pursue further inquiries or adopt new strategies in order to improve my epistemic situation?" Social *zetetic* prompting question: "Are the institutional or social group practices conducive to truth and understanding, or should they be altered in light of the present problem-situation?"	Improving the deliberative climate and nurturing the non-directly epistemic deliberative virtues in myself and others. Individual *zetetic* prompting question: "How can I acquire or apply deliberative virtues that will improve my own deliberations (and those of others in my community of inquiry)?" Social *zetetic* prompting question: "How should cognitive labor be divided in order to promote shared cognitive goals? Also, what part of that labor is it my responsibility to take on, and how do I responsibly assess the credibility of inquiry performed by others?"	Increasing my/our ethical awareness and considering what ethical aims and projects to pursue. Delineating ideals against which current practices and institutions may be criticized, and future practices designed. Individual *zetetic* prompting question: "What special responsibility for pursuing further inquiry, given the gravity of the situation on which my actions will bear, and the social roles I have (qua professional, parent, etc.)?" Social *zetetic* prompting question: "What socially 'transformative' aims should I endorse, and what individual, social and collective 'new' virtues should I posit as reflective of those aims?"

NO SUICIDE FOR PRESENTISTS: A RESPONSE TO HALES

Jimmy Alfonso LICON

ABSTRACT: Steven Hales constructs a novel argument against the possibility of presentist time travel called the *suicide machine argument*. Hales argues that if presentism were true, then time travel would result in the annihilation of the time traveler. But such a consequence is not time travel, therefore presentism cannot allow for the possibility of time travel. This paper argues that in order for the suicide machine argument to succeed, it must make (at least) one of two assumptions, each of which beg the question. The argument must either assume that the sequence of moments is invariant, or that time travel through time requires distinct, co-instantiated moments. Because the former disjunct assumes that presentist time travel is impossible and the latter assumes that presentism is impossible, the suicide machine argument fails.

KEYWORDS: presentism, suicide machine argument, time travel, Steven Hales

1. Introduction

Presentism is a theory of time that holds that only the present moment and those things that inhabit the present moment exist.[1] It is widely held that restricting the inventory of moments to the present moment creates a problem for the possibility of presentist time travel, since time travel seems to require moments other than the present and presentism denies such a possibility.[2] Appealing to the apparent tension between time travel and the presentist theory of time, Steven Hales[3] presents a novel argument against the possibility of presentist time travel called *the suicide machine argument*. The suicide machine argument holds that (i) time travel just is the process of an individual traveling between distinct, co-instantiated moments, and (ii) presentism denies that there are any moments other than the present to which an individual could travel. Thus given (i) and (ii), presentist time travel is a kind of suicide for the time traveler. I argue that the suicide machine argument

[1] See Ned Markosian, "A Defense of Presentism," in *Oxford Studies in Metaphysics*, vol. 1, ed. Dean Zimmerman (Oxford: Oxford University Press, 2003).

[2] Simon Keller and Michael Nelson, "Presentists Should Believe in Time-Travel," *Australasian Journal of Philosophy* 79 (2001): 333-45.

[3] Steven Hales, "No Time Travel for Presentists," *Logos & Episteme* I, 2 (2010): 353-360.

must either assume that the sequence of moments is invariant or that time travel presupposes the existence of two distinct, co-instantiated moments in order to succeed. Since both assumptions beg the question either against presentism or presentist time travel, the presentist is within her rights to reject the suicide machine argument outright. Furthermore, I provide a rough argument to demonstrate that presentist time travel is possible.

2. Proper Reconstruction

Hales maintains that presentist time travel would result in annihilation of the time traveler, rather than time traveling: "For presentists, getting into a time machine is suicide – the occupant goes out of existence. Recall that presentists are committed to a purely objective present; the events and objects at this objective present alone are real, even if other things have been or will be."[4] Hales thinks that presentism is committed to the suicide of time travelers because it is committed to the existence and objectivity constraints on the present moment. The *existence constraint* (EC) holds that concrete objects exist only if (i) they are capable of persisting through time and (ii) they are capable of occupying a particular moment. The *objectivity constraint* (OC) holds that the present moment just is the moment that exists to the exclusion of all other moments.

Hales supposes that the (OC) and (EC) constraints force the presentist to admit that her theory of time cannot account for time travel since there is no other moment besides the present for the time traveler to occupy, and as a result leaving the present would be equivalent to the suicide. Hales states: "Either presentists must identify the objective present with the present of external time or the present of Dr. Who's personal time. Suppose they [presentists] *identify the present of external time as the objective present* … In this case, by the Suicide Machine argument, Dr. Who leaves the objective present and thus goes out of existence."[5] One of the assumptions that results in a successful version of the suicide machine argument can explicated as: (PS) *the ordering of moments is invariant.*

Given the (PS) constraint, it would follow that although the passage of time is allowed under presentism, moments cannot occur out of sequence. There is a particular ordering imposed on a sequence of moments such that the ordering of moments is invariant. As such, constrain (PS) blocks the possibility of presentist time travel.

Hales continues: "Suppose instead presentists identify Dr. Who's personal time as the objective present. In this scenario, after he gets into the time machine

[4] Hales, "No Time Travel," 357.

[5] Hales, "No Time Travel," 358-9; *emphasis mine.*

everything else in the world is separated from the objective present, and hence the universe minus Dr. Who goes out of existence."[6] Hales holds that if the presentist rejects the objectivity of the present, then she is forced to accept either that the time traveler remains in existence to the exclusion of the rest of the universe, or posit a multiplicity of time lines.

> The suicide machine argument can be formulated as,
>
> S1: All objects O capable of persisting through time must occupy a particular moment (From [EC]).
>
> S2: The present moment exists to the exclusion of all other moments (From [OC]).
>
> So,
>
> S3: O must occupy the present moment (From [S1] and [S2]).
>
> S4: It is either the case that: (i) the ordering of moments is invariant, or (ii) there must be a moment that is distinct from the present moment to serve as an arrival moment for O.
>
> S5: But (i) prevents the possibility of presentist time travel and (ii) blocks the possibility of presentism.
>
> So,
>
> SC: O is not capable of presentist time travel (From [S3], [S4] and [S5]).

The disjunction (S4) applies specifically to presentist theories of time. It is constraints (EC), (OC) combined with disjunction (S4) that blocks the possibility of presentist time travel. The presentist cannot appeal to subjective considerations to save the possibility of presentist time travel given that she maintains the property of *being the present moment* holds of a moment subject independently. Appealing to subjective considerations is a viable move in an eternalist universe, rather than a presentist one.

3. Critical Comments

In spite of the initial plausibility of the suicide machine argument, the presentist need not forgo the possibility of time travel. The constraints upon which the suicide machine argument depends for its success beg the question either by assuming the impossibility of presentism generally, or presentist time travel specifically. Before exploring the ways in which the suicide machine argument begs the question, we must distinguish between presentism and eternalism as conflicting

[6] Hales, "No Time Travel," 359.

theories about the nature of time. Specifically, how it is that eternalism and presentism differ in their conception of the present.

Eternalism holds that all moments are metaphysically on par: past and future moments are just as real as present moments.[7] Eternalism explains the appearance of temporal privilege by appealing to the perspective of a given individual at a particular time, rather than an intrinsic property of the moment itself. Although eternalists are committed to objective temporal properties, (e.g. Event E_1 <occurred earlier than> E_2), they maintain that the present is privileged only perspectivally. For example, some eternalists might explain the apparent privilege of the present in epistemic terms. Those individuals that occupy moment M have this impression because of their unique epistemic access to the components of M. From the subject's point of view, the present moment appears privileged, but rather it is that the subject is epistemically privileged. Metaphysically, eternalism holds that no temporal moment is privileged above any other moment.

In contrast, presentism locates the difference-making properties of the present from the non-present moments in the moment itself. For example, some presentists might explain the difference-making properties of present moments in terms of actualization. The present moment and those objects and properties that inhabit the present moment are metaphysically privileged because they are actualized, while those objects that occupied past moments and will occupy the future moments are not privileged because they are not actualized. It is because presentism holds that the difference-making properties that differentiate the present moment from non-present moment are intrinsic to the moments themselves that they hold that the present is objectively privileged.[8]

> Further, all presentists accept a weak constraint on the objectivity of the present,
>
> (Exclusive Objectivity): T is the present moment *only if* T is the moment that exists to the exclusion of all other moments.

But the exclusive conception of the present alone is not sufficient to establish the suicide machine argument. It could be that the (objective) present moment changes, with the help of a time machine, to accommodate the time traveler. In such a scenario, a time machine would have the capacity to rearrange all of the matter and energy in the universe, such that it is indistinguishable from a past or future moment and the moment brought about by the time machine exists to the exclusion of all other moments. For example, it could be the case that a moment

[7] Hales, "No Time Travel," 358.

[8] Dean Zimmerman, "The A-theory of Time, the B-theory of Time, and 'Taking Tense Seriously'," Dialectica 59 (2005): 401-57.

from the Jurassic period is instantiated after a moment from the age of hovercrafts, *with the help of a presentist time machine.* To block such a move, Hales needs to appeal to a much stricter variety of temporal objectivity such as,

> (Fixed Objectivity): T is the present moment *only if* (i) T is the moment that exists to the exclusion of all other moments, and (ii) moments cannot occur out of sequence.

Presentists generally must endorse the exclusive objectivity of the present if they are to remain consistent presentists. In contrast, fixed objectivity is a far more stringent conception of the present. Even though appealing to the fixed conception of the present would result in a successful suicide machine argument, Hales has other options. It might be that Hales holds a disjunction akin to (S4): it is either the case that (i) the ordering of moments is invariant, or (ii) there must be a moment that is distinct from the present moment to serve as an arrival moment for O.

The problem with the suicide machine argument is that conclusion (SC) only follows from premises (S1) through (S5), if premise (S4) succeeds. Although a presentist can consistently affirm premise (S4) by affirming disjunct (i), there is no reason that the presentist must accept premise (S4) without further argumentation. In order for the suicide machine argument to succeed, it must appeal to (at least) one of the following constraints (each constraint serves as a disjunct that forms [S4]),

i. Temporal passage is fixed such that the ordering of moments is invariant.

ii. There must be distinct, co-instantiated moments to allow for the possibility of time travel.

It is not clear why the presentist would accept either conditions (i) or (ii) without further motivation. Presentism simpliciter is inconsistent with (ii), as such a condition rules out the persistence of objects from one moment to another, and thus begs the question against presentism. If distinct and co-instantiated moments were required to allow for the possibility of presentist time travel, then such a requirement would require distinct and co-instantiated moments to allow for the possibility of object persistence. Furthermore, disjunct (i) begs the question against the possibility of presentist time travel. If moments can only occur in an invariant ordering, and time travel allows for moments to occur in a different ordering, then time travel is not possible in a presentist universe. But this is a serious problem for the suicide machine argument, as it was supposed to provide justification for the claim that present time travel results in the suicide of the time travel. Since either condition (i) or (ii) is necessary for the success of the suicide machine argument,

then the suicide machine argument is guilty of begging the question against either presentism or presentist time travel.

A counter to the suicide machine argument can be formalized as follows,

> CA1: If the suicide machine argument succeeds, then the presentist ought to hold either disjunct (i) or (ii).
>
> CA2: Disjunct (i) begs the question against presentist time travel.
>
> CA3: Disjunct (ii) begs the question against presentism.
>
> CA4: But the presentist ought not to hold either disjunct (i) or (ii) (From [CA2] and [CA3]).
>
> So,
>
> CA': The suicide machine argument does not succeed (From [CA1] and [CA4]).

The possibility of presentist time travel does not entail that the time traveler ceases to exist, but rather that it is possible for the present moment to accommodate the time traveler. Of course such a process must preserve the objective temporal exclusivity of the present.

For example, consider a fighter jet F that is confined to taking off from and landing on a particular aircraft carrier C. It does not follow that if F leaves C, then F will have no place to land. It just means that F must take off from and land on C. Suppose that F needs to travel several miles away from its current position. The landing constraint does not entail that F cannot land, but rather that: *if F can land in in a location other than the location that F took off from, then C must have moved from the departure to the arrival location.*

4. Positive Proposal

Given the failures of the suicide machine argument, we are in a position to construct an argument for the possibility of presentist time travel. Indeed, given a few stipulations presentist time travel is possible.

Consider the following argument,

> E1: If it is possible that: (a) object O can persist from moment M_1 to moment M_2, (b) it need not be the case that M_1 and M_2 are distinct and co-instantiated, and (c) the moment occupied by O just is the moment that exists to the exclusion of all other moments.
>
> Then presentist time travel is possible.
>
> E2: It is possible that (a), (b) and (c).
>
> So,

EC: Presentist time travel is possible.

Of course, one might wonder how the presentist could explain time travel, given that the present moment must transform to provide an arrival moment for the time traveler. All the presentist needs is the possibility of a time machine capable of operating in a presentist universe, transforming the present moment by annihilating the components of the departure moment and instantiating components of the arrival moment. The time machine constraint can be stated roughly,

> (Time Machine Constraint):[9] Time machine B is capable of an instance of presentist time travel only if B is capable of annihilating the departure moment and instantiating the arrival moment such that both the departure and arrival moments exist to the exclusion of all other moments.

The time machine constraint provides the presentist a rough response to the problem of presentist time travel mentioned by Hales: "Not only does presentist 'time travel' merely require the would-be traveler to go out of existence *in nihilum*, but it also requires that objects come into existence *ex nihilo*."[10] Hales claims that presentists must deal both with the conclusion of the suicide machine argument and account for the ex nihilo genesis of objects in the moment to which an individual traveled.

However, Hales' point is subject to a dilemma:

> For any given instance of presentist time travel, it cannot *both* be the case that: (a) the time traveler goes out of existence; and, (b) the arrival moment comes into existence *ex nihilo.*

If option (a) were the case, and the time traveler ceased to exist for lack of an arrival moment, then the *ex nihilo* genesis of an arrival moment would not be a problem. However, if we suppose there *were* an arrival moment for the time traveler to occupy, then the time traveler would not cease to exist for want of an arrival moment. Hales cannot consistently maintain that a particular instance of presentist time travel is subject to both problems (a) and (b), because both horns of the dilemma cannot be the case for any single instance of attempted time travel. Hales might respond that he only meant that either option (a) or (b) would be a problem for present time travel. But this response fails for another reason. Even individually, options (a) and (b) are only problematic for the presentist *if* she cannot account for the possibility of time travel or a possible mechanism for annihilating the departure moment and instantiating the arrival moment.

[9] For the purposes of this paper, I will assume toy physics.
[10] Hales, "No Time Travel," 358.

The primary purpose of proposing a possible presentist friendly time machine is to provide a rough response to problem (b), since we already have a response to problem (a). The arrival moment is not an instance of *ex nihilo* genesis, but rather the arrival moment is brought about by the time machine. It is plausible to think that presentist time travel is possible *only if* it is possible that there could be time machines capable of annihilating the departure moment and instantiating the arrival moment. A presupposition of something originating *ex nihilo* is that there is no causal mechanism capable of bringing that something about, but rather that it simply *comes from nothing*. However, a time machine capable of annihilating the departure moment and instantiating the arrival moment would block the charge that the arrival moment did came into existence *ex nihilo*.

A time machine could accomplish the annihilation the departure moment and the instantiation the arrival moment simply by rearranging all the matter and energy in the universe, so as to make it indistinguishable from a past or future moment. The possibility of a time machine can be easily generated,

> (ME)[11] If machine F is capable of rearranging all of the matter and energy in the universe, such that it is indistinguishable from a past or future moment, and the identity of indiscernibles[12] holds between the instantiated moment and the past or future moment, then F is a time machine.

The presentist is entitled to stipulate the possibility of a machine that meets the antecedent of conditional (ME), and as such she is entitled to postulate the possibility of a presentist time machine. Since we are only concerned with bare possibilities, it is the advocate of the suicide machine argument that must provide the argument deriving a contradiction from the assumptions made on behalf of presentist time travel.

Now that have disposed of condition (i)[13] via the possibility of a presentist-friendly time machine, is there a compelling reason for the presentist to take condition (ii)[14] seriously? It might be argued that in order for a time machine to annihilate the departure moment while instantiating the arrival moment would require *both the destruction and construction of the same moment*. But then how could it be the case that the time machine is *both* annihilating and instantiating the present moment? The presentist is subject to following prima facie trilemma,

---

[11] I am indebted to Bernard Molyneux (Philosophy Department, U.C. Davis) for this insight.

[12] If, for every property F, object x has F *if and only if* object y has F, then x is identical to y.

[13] i. Time is fixed such that the ordering of moments is invariant.

[14] ii. There must be distinct, co-instantiated moments so as to allow for the possibility of presentist time travel.

> Either (i') a time machine can both annihilate and instantiate the same moment, (ii') temporal moments can be indeterminate, or (iii') there can be distinct, co-instantiated moments.

The presentist must outright reject (iii') as it conflict with presentism. But how might the presentist evaluate options (i') and (ii')? She cannot accept option (ii') for fear of inconsistency. If she accepts the possibility of indeterminate moments (i.e. moments that lack facts of the matter), then she must explain how it is possible for a time machine to occupy an indeterminate moment. This is difficult since 'a time machine occupying moment M' would be a fact of the matter. The presentist is committed to the claim that every existent that is capable of persisting through time must inhabit the present moment. But if the present moment were indeterminate, then it is unclear how a time machine could occupy the present moment.

The critic holds that the presentist is not in any better position with regard to option (i'). If the presentist accepted option (i'), then she is forced to admit that the same moment can be in the process of being annihilated and instantiated. Since such a claim cannot be parsed in temporal terms, the critic might argue that an individual moment must possess both annihilation and instantiation processes. Why should the presentist accept option (i')? It would seem that option (i') presupposes a principle that the presentist is free to reject: an individual moment cannot co-possess annihilation and instantiation processes. But such a presupposition requires a further presupposition to block the presentists rejection of option (i'): *there cannot be transitional moments.*

A transitional moment can be defined as,

> (TM) Moment M is a transitional moment only if (a') M exists to the exclusion of all other moments, and (b') M contains components of both a departure moment and an arrival moment.

A transitional moment just is a moment that exists to the exclusion of all other moments in the process of having all of its matter and energy rearranged by a time machine. Transitional moments allow the presentist to explain how it can be the case that a time machine can annihilate the departure moment and instantiate the arrival moment without postulating distinct, co-instantiated moments. Any time machine capable of traveling through time in a presentist universe must be capable of generating transitional moments.

5. Conclusion

The suicide machine argument holds that if presentism were true, then an individual attempting to travel through time would cease to exist. This is because

presentism holds that the present moment exists to the exclusion of all other moments and time travel seems to require moments that are distinct from the present. Hales holds that such constraints force the presentist to admit that time travel in a presentist universe is equivalent to going out of existence.

I have argued that this conclusion is hasty. There is no reason for the presentist to deny that a time machine could simply accommodate the time traveling of an individual by rearranging all the matter and energy in the universe at particular moment such that it is indistinguishable from a past or future moment. It is possible for an inhabitant of a presentist universe to travel through time if there is a time machine capable of accommodating her time traveling. A critic might object that time travel is possible only if there are two distinct, co-instantiated moments such that a time machine can annihilate aspects of the departure moment and instantiate aspects of the arrival moment. But such a response presupposes that transitional moments are not possible.

The considerations offered in favor of the possibility of presentist time travel are modest at best. I simply maintain that the suicide machine argument fails to block the possibility of presentist time travel, and a modest conditional: *If the suicide machine argument is the only argument against the possibility of presentist time travel, then presentist time travel is possible.*

REVIEWS

Scott Aikin, *Epistemology and the Regress Problem* (New York, London: Routledge, 2010)

Reviewed by Eugen Huzum*

Definitely, *Epistemology and the Regress Problem* deserves a special and close attention, at least from those interested in topics such as epistemic justification, argumentation, rationality, and the like. The book has at least five qualities that recommend such attention.

First: *Epistemology and the Regress Problem* is very intriguing, challenging and thought provoking (even for experienced and well-trained professional philosophers). This is because, although it offers other things too, the book is, above all, a (very good, careful and ingenious) defense of epistemic infinitism. But, as Aikin himself highlights, infinitism – which is, essentially, the view that epistemic justification requires an infinite chain of reasons (or, in other words, the view that, in order to have a justified belief, we must have an infinite series of nonrepeating, supporting reasons) – is a strongly counterintuitive view, being considered 'obviously false,' 'patently absurd' or even 'crazy' by almost anyone, philosophers (foundationalists, coherentists, contextualists or foundherentists) included. In addition, for the overwhelming majority of contemporary or past epistemologists, infinitism is also a 'dangerous' doctrine, because of its supposedly inevitable skeptical implications about epistemic justification (and, if we accept that justification is a necessary condition for knowledge, about human knowledge too).

Being a defender of infinitism, Aikin, of course, disagrees. In his (again, to my mind, well argued) opinion, on close reflection, infinitism (or, more exactly, a particular version of infinitism) proves to be the correct view of epistemic justification (or at least a view that is not at all so implausible as it appears *prima facie*). Yes, Aikin recognizes, infinitism may entail skepticism. But this is not a decisive argument against it. The fact that infinitism may entail skepticism does not prove that infinitism is a false view about the requirements for epistemic justification and knowledge. Skepticism or epistemic pessimism may be, after all, the correct view about the possibility of human knowledge and epistemic justification.

* ACKNOWLEDGEMENT: This paper was made within *The Knowledge Based Society Project* supported by the Sectorial Operational Program Human Resources Development (SOP HRD), Financed by the European Social Fund and by the Romanian Government under the contract POSDRU/89/1.5/S/56815.

Second: thanks to its author's writing style – which is very clear, precise, friendly, not excessively technical and often humorous – *Epistemology and the Regress Problem* is not only a provocative, but also a very pleasant and enjoyable reading. This feature makes me strongly recommend the book not only to professional epistemologists, but also to students or even to those for whom philosophy/epistemology is only a collateral concern or simply a hobby.

Third: the book offers a very interesting and, to my knowledge, a quite new interpretation or 'redescription' of its main topic: the so-called problem (or, as it is often considered, the 'danger') of the infinite regress of epistemic justification/argumentation. As interpreted by Aikin (pp. 22-32), the regress problem is a paradox. The paradox is constituted, *in extenso*, by the tension/contradiction between: 1) the principle of inferential justification ("If S is justified in holding p, then there is some q that S is justified in holding, and q supports p for S"), 2) the principle of transitivity of inferential support ("If S's commitment to p is supported by S's commitment to q and S's commitment to q is supported by S's commitment to s, the S's commitment to p is supported by S's commitment to s"), 3) the principle of asymmetry of inferential support ("In order for S to be justified in holding that p, even if p and another of S's commitments, q, are mutually supporting, p must have independent support from some other commitment, r, that S is justified in holding"), which taken together imply infinitism as a requirement of epistemic justification, and 4) the principle of finitism ("If S is justified in holding that p, then S's chain of supporting reasons, C, for p is finite"). As Aikin summarizes his point, "the core of the regress problem is that it seems that supporting reasons need to come in infinite chains, but there cannot be any such chains." (p. 34)

Fourth: As its author also highlights, *Epistemology and the Regress Problem* is the first book-length defense of epistemic infinitism as a solution to the infinite regress problem (the solution being, obviously, that of arguing that the principle of finitism is false). More importantly, the book is the most 'updated' and complete defense of epistemic infinitism. All major anti-infinitist arguments find a rejoinder here (pp. 50-71). At the same time, the book defends with great depth all fundamental premises and commitments that lead to infinitism: 'epistemic aspirationalism,' responsibilism and evidentialism, 'anti-populism' in knowledge/justification-attribution, 'doxastic ascent' or 'meta-justification' (pp. 3-49), and, above all, antidogmatism (pp. 158-179).

Fifth: Aikin's theory of epistemic justification is a new version of infinitism. In other words, his infinitism is different from that of Charles Sanders Peirce, Peter Klein or Jeremy Fantl. As Aikin calls it, his theory is a "*strong,*

impure, synchronic, emergent infinitism." (p. 79, Aikin's emphasis.) It is 'strong' because it holds that infinitely extended chains of inference are necessary for any justification tree, it is 'impure' because it allows more sources or forms of epistemic support than infinitely extended chains of inference (accepting and defending, for example, the foundationalist idea that 'basic' or independent reasons are also necessary components for a justificatory story), it is 'synchronic' because it says that a subject S is justified in holding a belief p at *a time t*, if S has an infinitely long supporting chain of reasons for p *at t*, and finally, it is 'emergent' because it holds that the support provided for a belief by the infinite series of reasons *emerges from* the fact that the commitment belongs to the series, not that it is *transmitted through* the infinite series. The main feature that distinguishes Aikin's infinitism is its impurity/consistency with foundationalism. This is also the most important feature of his theory, since, as Aikin shows (pp. 102-111), it is a feature which helps him to solve one of the most powerful arguments against infinitism: the '*modus ponens reductio*' of it (the argument that on the infinitist theory of epistemic justification, we have no rational way of telling the difference between one justifying set of beliefs that is conducive of truth and one that is not).

I am convinced that, if it will receive the attention it deserves, *Epistemology and the Regress Problem* will have a significant impact in the theory of epistemic justification/rationality. Of course, I do not expect it to convince all epistemologists (or at least most of them) that "strong, impure, synchronic, emergent infinitism" is indeed the best theory about the structure of epistemic justification/rationality. Such an expectation is undoubtedly unrealistic (due to the highly divergent deep presuppositions, convictions and commitments of the philosophers who are striving to provide a theory of epistemic justification). But I believe that the book will have at least two important effects. The first is that, due to the high standards of its argumentation, it will further improve the image of infinitism, at least among professional epistemologists. Actually, to improve the prospects of infinitism and to show that it deserves at least a place at the table in philosophical discussions of epistemic justification is the main objective of its author. The second – and, perhaps, more important – effect of the book will be that of enhancing the philosophical debates on its main topic, since it will compel those epistemologists who will remain unconvinced by Aikin's arguments to revise, improve, refine or even renew their arguments against infinitism and for alternative theories of epistemic justification/rationality.

NOTES ON THE CONTRIBUTORS

Scott F. Aikin is Senior Lecturer at the Philosophy Department of Vanderbilt University. His areas of interest are: epistemology, informal logic, argumentation, pragmatism, and ancient philosophy. His books include: *Epistemology and the Regress Problem* (2011, reviewed in this issue of *Logos & Episteme*), *Reasonable Atheism* (2011, with Robert B. Talisse), and *Pragmatism: A Guide for the Perplexed* (2008, with Robert B. Talisse). Contact: scott.f.aikin@vanderbilt.edu.

Guy Axtell is Assistant Professor in the Philosophy and Religious Studies Department, Radford University. His research areas are epistemology and analytic and comparative philosophy of religion. His most recent publications include: "Three Independent Factors in Epistemology" (*Contemporary Pragmatism*, 2010, with Phillip Olson), "Character-Trait Ascription in Ethics and Epistemology" (*Metaphilosophy*, 2010), and "Virtue Theoretic Responses to Skepticism" (*Oxford Handbook of Epistemology,* John Greco, ed., 2009). Contact: gsaxtell@radford.edu.

Cătălin Bobb has a Ph.D. in philosophy at Babeş Bolyai University, Romania. His main interests are hermeneutics, phenomenology and philosophy of religion. He is currently working as a postdoctoral fellow at the Romanian Academy, in a project regarding the problem of explication and understanding. Contact: catalinbobb@yahoo.com.

Michael Harbour is a Ph.D. candidate in the Department of Philosophy at Vanderbilt University with a thesis on *Freedom as Anti-Domination*. His research focuses on social and political philosophy. He is the author of "Nagel on Public Education and Intelligent Design" (*Journal of Philosophical Research*, 2010, with Scott F. Aikin and Robert B. Talisse) and "The Ethics of Inquiry and Engagement: The Case of Science in Public" (*Public Affairs Quarterly*, 2010, with Scott F. Aikin).

Jimmy Alfonso Licon is M.A. Candidate in Philosophy, Department of Philosophy, San Francisco State University. His main areas of philosophical interest are: epistemology (responses to skepticism, epistemic theories of justification, introspective skepticism, and moral epistemology), metaphysics and philosophical methodology. Contact: jalicon@mail.sfsu.edu

Ned Markosian is Professor in the Department of Philosophy at Western Washington University. His area of specialization is metaphysics and his areas of competence include epistemology, ethics, logic, modern philosophy, ancient philosophy, philosophy of language, philosophy of religion, philosophy of science. He is the author of *An Introduction to Metaphysics* (with John W. Carroll, 2010). Contact: ned.markosian@wwu.edu

Jonathan Matheson is Assistant Professor of Philosophy at the University of North Florida. His areas of specialization are epistemology and ethics, and his areas of competence are metaphysics, philosophy of language, and philosophy of religion. His most recent work includes: "Epistemological Considerations Concerning Skeptical Theism: A Response to Dougherty" (*Faith and Philosophy*, forthcoming), "Bergmann's Dilemma: Exit Strategies for Internalists" (*Philosophical Studies*, 2011, with Jason Rogers), "Conciliatory Views of Disagreement and Higher-Order Evidence" (*Episteme: A Journal of Social Philosophy*, 2009). Contact: j.matheson@unf.edu

Christian Möckel is Professor of Philosophy at Humboldt University (Berlin) and University of Hamburg. His academic interests include Ernst Cassirer's work, Husserl's phenomenology, political philosophy of the 17th century and the relationship between Life and Reason in its multiple philosophical aspects. His books include *Das Urphänomen des Leben. Der Lebensbegriff Ernst Cassirers* (2005), *Phänomenologie. Probleme, Bezüge, Interpretationen* (2003), and *Einführung in die transzendentale Phänomenologie* (1998). Contact: MoeckelC@Philosophie.hu-berlin.de

Jonathan Neufeld is Assistant Professor of Philosophy at the College of Charleston. His research interests are in: philosophy of music, aesthetics, political philosophy, and philosophy of law. He is particularly interested in problems surrounding performance and interpretation. He is currently working on two books: *Music in Public: How Performance Shapes Democracy* (forthcoming 2011), *and Listeners, Critics, and Judges* (in preparation). Contact: neufeldja@cofc.edu

Robert B. Talisse is Professor of Philosophy and Political Science and Director of Graduate Studies in Philosophy at Vanderbilt University. He is also editor of *Public Affairs Quarterly*. His area of specialization is contemporary political philosophy. He is the author of: *Reasonable Atheism* (2011, with Scott Aikin), *Democracy and Moral Conflict* (2009), *Pragmatism: A Guide for the Perplexed* (2008, with Scott

Aikin), *A Pragmatist Philosophy of Democracy* (2007), and *Democracy After Liberalism* (2005). Contact: robert.talisse@vanderbilt.edu.

Hamid Vahid is Professor of Philosophy and the Head of the Analytic Philosophy Faculty at the Institute for Fundamental Sciences in Tehran, Iran. He is the author of *Epistemic Justification and the Skeptical Challenge* (2005) and *The Epistemology of Belief* (2008). His list of articles include: "Rationalizing Beliefs: Evidential vs. Pragmatic Reasons" (*Synthese*, forthcoming), "The Concept of Entitlement and its Epistemic Relevance" (*Pacific Philosophical Quarterly*, 2011), and "Alston on Belief and Acceptance in Religious Faith" (*Heythrop Journal*, 2009). Contact: vahid@ipm.ir.

Mark Owen Webb is Associated Professor in the Department of Philosophy, Texas Tech University. His areas of specialization are epistemology and philosophy of religion and his areas of competence are Asian philosophy, Buddhism and feminism. He is currently working on a book entitled *A Comparative Doxastic-Practice Epistemology of Religious Experience* (under contract with Springer Press). He also edited, with Heidi Grasswick, *Feminist Epistemology as Social Epistemology* (a special issue of *Social Epistemology* 16, 2002). Contact: Mark.Webb@ttu.edu.

LOGOS & EPISTEME:
AIMS & SCOPE

Logos & Episteme is a quarterly open-access international journal of epistemology that appears in printed and on-line version in March, June, September, and December. Its fundamental mission is to support and publish various current reflections and researches that aim at investigating, analyzing, interpreting or philosophically explaining the human knowledge in all its aspects, forms, types, dimensions or practices.

For this purpose, the journal will publish articles, reviews or discussion notes focused as well on problems concerning the general theory of knowledge, as on problems specific to the philosophy, methodology and ethics of science, philosophical logic, metaphilosophy, moral epistemology, epistemology of art, epistemology of religion, social or political epistemology, epistemology of communication. Studies in the history of science and of the philosophy of knowledge, or studies in the sociology of knowledge, cognitive psychology, and cognitive science are also welcome.

The journal intends to promote all methods, perspectives and traditions in the philosophical analysis of knowledge, from the normative to the naturalistic and experimental, and from the Anglo-American to the continental or Eastern.

The journal accepts for publication texts in English, French and German, which satisfy the norms of clarity and rigour in exposition and argumentation.

NOTES TO CONTRIBUTORS

1. Accepted Papers

The journal accepts for publication articles, discussion notes and book reviews submitted exclusively to *Logos & Episteme* and not published, in whole or substantial part, elsewhere. The editors of *Logos & Episteme* reserve the right to refuse all future papers belonging to the authors who fail to observe this rule.

2. Submission Address

Please submit your manuscripts electronically at: logosandepisteme@yahoo.com. Authors will receive an e-mail confirming the submission. All subsequent correspondence with the authors will be carried via e-mail. When a paper is co-written, only one author should be identified as the corresponding author.

3. Paper Size

The articles should normally not exceed 12000 words in length, including footnotes and references. Articles exceeding 12000 words will be accepted only occasionally and upon a reasonable justification from their authors. The discussion notes must be no longer than 3000 words and the book reviews must not exceed 4000 words, including footnotes and references. The editors reserve the right to ask the authors to shorten their texts when necessary.

4. Manuscript Format

Manuscripts should be formatted in Rich Text Format file (*rtf) or Microsoft Word document (*docx) and must be double-spaced, including quotes and footnotes, in 12 point Times New Roman font. Where manuscripts contain special symbols, characters and diagrams, the authors are advised to also submit their paper in PDF format. Each page must be numbered and footnotes should be numbered consecutively in the main body of the text and appear at footer of page. For all references authors must use the Humanities style, as it is presented in The Chicago Manual of Style, 15th edition. Large quotations should be set off clearly, by indenting the left margin of the manuscript or by using a smaller font size. Double

quotation marks should be used for direct quotations and single quotation marks should be used for quotations within quotations and for words or phrases used in a special sense.

5. Official Languages

The official languages of the journal are: English, French and German. Authors who submit papers not written in their native language are advised to have the article checked for style and grammar by a native speaker. Articles which are not linguistically acceptable may be rejected.

6. Abstract

All submitted articles must have a short abstract not exceeding 200 words in English and 3 to 6 keywords. The abstract must not contain any undefined abbreviations or unspecified references. Authors are asked to compile their manuscripts in the following order: title; abstract; keywords; main text; appendices (as appropriate); references.

7. Author's CV

A short CV including the author`s affiliation and professional address must be sent in a separate file. All special acknowledgements on behalf of the authors must not appear in the submitted text and should be sent in the separate file. When the manuscript is accepted for publication in the journal, the special acknowledgement will be included in a footnote on the first page of the paper.

8. Review Process

The reason for these requests is that all articles, with the exception of articles from the invited contributors, will be subject to a strict blind-review process. Therefore the authors should avoid in their manuscripts any mention to their previous work or use an impersonal or neutral form when referring to it. The review process is intended to take no more than six months. Authors not receiving any answer during the mentioned period are kindly asked to get in contact with the editors. Processing of papers in languages other than English may take longer. The authors will be notified by the editors via e-mail about the acceptance or rejection of their papers. The editors reserve their right to ask the authors to revise their papers and

the right to require reformatting of accepted manuscripts if they do not meet the norms of the journal.

9. Acceptance of the Papers

The editorial committee has the final decision on the acceptance of the papers. Papers accepted will be published, as far as possible, in the order in which they are received and they will appear in the journal in the alphabetical order of their authors.

10. Responsibilities

Authors bear full responsibility for the contents of their own contributions. The opinions expressed in the texts published do not necessarily express the views of the editors. It is the responsibility of the author to obtain written permission for quotations from unpublished material, or for all quotations that exceed the limits provided in the copyright regulations. The papers containing racist and sexist opinions assumed by their authors will be rejected. The presence in texts of sexist or racist terms is accepted only if they occur in quotations or as examples.

11. Checking Proofs

Authors should retain a copy of their paper against which to check proofs. The final proofs will be sent to the corresponding author in PDF format. The author must send an answer within 3 days. Only minor corrections are accepted and should be sent in a separate file as an e-mail attachment.

12. Reviews

Authors who wish to have their books reviewed in the journal should send them at the following address: Institutul de Cercetări Economice şi Sociale „Gh. Zane" Academia Română, Filiala Iaşi, Str. Teodor Codrescu, Nr. 2, 700481, Iaşi, România. The authors of the books are asked to give a valid e-mail address where they will be notified concerning the publishing of a review of their book in our journal. The editors do not guarantee that all the books sent will be reviewed in the journal. The books sent for reviews will not be returned.

13. Property & Royalties

Articles accepted for publication will become the property of *Logos & Episteme* and may not be reprinted or translated without the previous notification to the editors. No manuscripts will be returned to their authors. The journal does not pay royalties. Authors of accepted manuscripts will receive free of charge two copies of the issue containing their papers.

14. Permissions

Authors have the right to use their papers in whole and in part for non-commercial purposes. They do not need to ask permission to re-publish their papers but they are kindly asked to inform the Editorial Board of their intention and to provide acknowledgement of the original publication in *Logos & Episteme*, including the title of the article, the journal name, volume, issue number, page number and year of publication. All articles are free for anybody to read and download. They can also be distributed, copied and transmitted on the web, but only for non-commercial purposes, and provided that the journal copyright is acknowledged.

Bun de tipar: 2011 • Apărut: 2011 • Format 17 × 24 cm

Iaşi, str. Grigore Ghica Vodă nr. 13
Tel. Dif.: 0788.319462; Fax: 0232/230197
editura_ie@yahoo.com; www.euroinst.ro