Volume 1 Issue 1 2014

Symposion

Theoretical and Applied Inquiries in Philosophy and Social Sciences

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

Contact address:
Institutul de Cercetări Economice și Sociale „Gh.Zane"
Iași, str. T.Codrescu, nr.2, cod 700481
Tel/ Fax: 004 0332 408922
Email: symposion.journal@yahoo.com
www.symposion.acadiasi.ro/

ISSN: 1584-174X

This journal was edited within "Innovation and Development in the Patterning and Representation of Knowledge Through PhD and Post-PhD Fellowships" 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 159/1.5/S/133675.

Invest in People

The content of these materials does not necessarily represent the official position of the European Union and of the Romanian Government.

TABLE OF CONTENTS

Introductory Note

Teodor Dima

With the present issue of *Symposion: Theoretical and Applied Inquiries in Philosophy and Social Sciences*, we continue and improve an itinerary started in 2003 with the Romanian journal *Symposion – Revistă de ştiinţe socio-umane* (*Symposion – A Journal of Humanities*) by adding new roadways and fortifying its ground by enlarging the team of remarkable scholars and editors that will work in the service of its mission.

Now, the new *Symposion* goes beyond Romania's borders, heading towards the four cardinal points of the globe. As its predecessor, the journal opens its pages to different fields of philosophical and scientific investigation: from aesthetics, philosophy of art, epistemology and philosophy of science, logic, ethics, hermeneutics, rhetoric, history of philosophy, metaphysics, metaphilosophy, philosophy of religion, philosophy of mind to cognitive science, anthropology, history and cultural studies, communication and media studies, economics, law, psychology, sociology, education, social work, and political sciences.

Although the first issue gathers papers mainly from philosophy, *Symposion* wishes to offer researchers in all the above-mentioned fields a new environment, favourable to original contributions, discussions, debates, and critical opinions.

The editorial board is committed to continue to promote Romanian contributions in humanities and social sciences and also to continuously improve *Symposion*, making it attractive to as many philosophers and social and political scientists as possible and observing the highest quality standards for philosophical and scientific research.

The idea of opening *Symposion* to contributions from all over the world and from different areas of research emerged a few years ago, when our Institute issued *Logos & Episteme. An International Journal of Epistemology*, which, as its title suggests, was primarily dedicated to publishing inquiries in the field of epistemology. The editorial board noticed that, due to this limitation, many interesting, original and well-written contributions submitted to *Logos & Episteme* had to be rejected on the ground of not fitting in the scope of the journal. This was one of the main reasons of transforming the old *Symposion* in a new international journal destined both to welcome contributions in a larger

Symposion, 1, 1 (2014): 5-6

area of philosophical and scientific investigation and to respond in a better way to our Institute's wider range of research interests.

The second main reason of *Symposion*'s transformation was the enthousiastic feed-back that the editorial board of *Logos & Episteme* received from the philosophical community since its publication. Therefore, some of the members of the editorial board of *Logos & Episteme* that are working in our Institute decided to create a new *Symposion*, with the professed hope that it will pursue its quest for success „non numero, sed pondere" (Cicero, *De officiis*, II, 2).

Research Articles

Stoicism, Feminism and Autonomy

Scott Aikin and Emily McGill-Rutherford

Abstract: The ancient Stoics had an uneven track record with regard to women's standing. On the one hand, they recognized women as fully capable of rationality and virtue. On the other hand, they continued to hold that women's roles were in the home. These views are consistent, given Stoic value theory, but are unacceptable on liberal feminist grounds. Stoic value theory, given different emphasis on the ethical role of choice, is shown to be capable of satisfying the liberal feminist requirement that autonomy must be respected. In turn, a model for Stoic feminism is proposed.

Keywords: stoicism, feminism, autonomy, liberalism, Epictetus, Seneca

I

That the Stoics had proto- or incomplete feminist commitments is a relative commonplace in the critical literature on Stoicism.[1] On the one hand, the Stoics thought that women were equals with men in their standing as rational beings.[2] On the other hand, the Stoics, despite their progressive views in principle, were socially conservative in practice. Women may have had equal capacities for virtue as men, but they nevertheless had different natural and social roles to play. So the Stoics held that women were to be offered different opportunities.[3]

[1] Additionally, there is wide agreement that the lacuna of explicit Stoic work on the tension is troubling. See: Arnold 1911, Manning 1973, Pomeroy 1976, Asmis 1996, Hill 2001, Nussbaum 2002, Engel 2003, and Bates 2014.

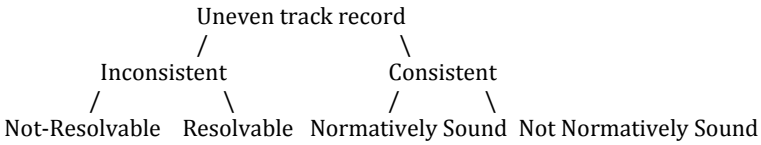
[2] Zeno's Republic has men and women sharing equal standing (DL VI.12). Cleanthes holds that men and women are equal in virtue (DL VII.175). Musonius Rufus holds that women should be taught philosophy (Stobaeus. 2.31.126). Seneca holds that women have the same capacities as men for virtue (*Cons. Marc*. 16.1). Epictetus argues that women are equal by nature (D.3.22.68). The stoics traced their philosophical lineage through the Cynics to Socrates, who all held similarly progressive views on women's equal capacities. See Socrates's proposal in Plato's *Republic* is that women can serve equally as guardians and philosopher kings in the good city (*Rep* V.451d). Diogenes and Antisthenes hold that women can philosophize (DL VII.12), and the Cynic Crates also had an equally philosophically adept wife, Hipparchia (DL VI.96-8).

[3] Like Socrates' views on women guardians, Zeno's early views on liberty were more for minimizing social strife than for the sake of women's liberation. Similarly, Musonius holds that women should learn philosophy, because such training would make them better (wiser and more dutiful) housewives (Stobaeus 2.31.127). Seneca, despite holding that women have the same native capacity for virtue, nevertheless also holds that there are special impediments to virtue that come with being a woman: lack of self-control (*Ad Helv*. 14.2), credulity (*De Cons*.

Let us call this phenomenon of these two conflicting trends *Stoicism's uneven track record*.

In light of the uneven track record, a constellation of interpretive and evaluative questions arise. These questions come in two orders. The first-order questions are strictly interpretive: What is the relationship between the currents of Stoic progressivism and Stoic misogyny? Are they consistent? The second-order questions regard the first-order answers, and they are mostly evaluative. If the currents of Stoic thought are inconsistent, which is the better (both for Stoic consistency and for normative soundness) view to jettison? If the two trends are consistent, given broader Stoic value theory, are the results normatively sound?

As a consequence of the orders of questions, we have a relatively simple taxonomy for interpretive takes on Stoicism's uneven track record.



Our argument will proceed in two stages. First, we will argue that the Stoics' progressive view about women's capacities is consistent with their conservative views about women's roles, but this consistency is a morally unsound consistency. We will hang our case for consistency on Epictetus' *Enciridion* 40, which simultaneously manifests both trends of the uneven track record and highlights the unacceptable elements of the Stoic program with women.

The second stage of our argument is that though the Stoics' uneven track record is internally consistent but morally wrong, it did not *have to* be that way. Stoicism's value theory provides sufficient material for not only the in-principle argument that women are equal partners with men in the *cosmopolis* (or as citizens of the world), but that unequal opportunity is unjust, misogyny is a failure to recognize the dignity of a fellow rational creature, and individual choice is a deep source of moral value. That the Stoics had an uneven track record by feminist lights is not evidence that Stoicism *must have* such a problem. As a consequence, a consistent and morally sound Stoic feminism is possible.

Sap. 19.2), and simple-mindedness (*Ad Marc*. 16.3). And Epictetus standardly references women as the kind of humans who can't keep their emotions in check (D 3.24.53) or as the kind of pretty trophy one would want when living the life of externals (D 4.94). This is not to mention all the standard usages of casually misogynistic phraseology. "Philosophize like a man, don't simper like a woman" (Seneca: *De Const.* I.1.2).

II

We will begin with a focused reading of Epictetus' *Enchiridion* 40. Our objective is to show the overlap of the two trends of progressivism and misogyny. We'll then turn to asking whether this is a necessary connection for the Stoics.

Epictetus' E40 is addressed to young men, Stoic progressors. It functions on two levels. First, it is an exercise in culture criticism: Epictetus observes and condemns the sexualizing of young women and the way they internalize this way of viewing themselves. Second, it is a call for action – the men in these women's lives must not only not participate in this activity, but call young women's attention to it and offer them an alternative of modesty and uprightness.

> Women from fourteen years old are flattered with the title of 'mistresses' by men. Therefore, perceiving that they are regarded only as qualified to give the men pleasure, they begin to adorn themselves, and in that to place all their hopes. We should, therefore, fix our attention on making them see that they are valued solely for displaying decent, modest and discreet behavior. (E40)

Here is how E40 manifests the two trends. On the one hand, the cultural critical element not only is a focus on the norms of objectifying women but it also acknowledges that women have the capacity to recognize this cultural pressure and can refuse to participate in it. This is the progressive element: many cultural norms fail to recognize the dignity of women as rational creatures, and women who participate in these norms begin to lose sight of that dignity, too. On the other hand, there is the alternative posed, that of modesty and uprightness. The problem is that the alternative posed is not altogether much better than the thought criticized. Why is being demure the *only* way one is honored?

In the same way that Epictetus' teacher, Musonius Rufus, held that philosophy for women yields wiser and more dutiful housewives, Epictetus seems to think that philosophy for daughters is for yielding modest young women. The trouble is that these trends, that of criticizing repressive cultural norms for the sake of encouraging other exclusions, seem inconsistent. This, again, is the unevenness of the Stoics' track record. Epictetus criticizes a cultural norm that offers young women only one role for them to play – that of a sexual object. He then offers an alternative – that of demure modesty. But this alternative, again, is only but one role, and it is not one that respects the variety of forms of human dignity. It is, again, merely a role to play. Epictetus may be right that sex-object is not a role that expresses a woman's dignity (and certainly hanging all one's hopes on it doesn't), but the better criticism should be not with *that option*, but with its *exclusivity* – that it is the *only option* for self-worth. Again, Epictetus, then, poses an exclusive option. On the face of it, Epictetus' views are inconsistent.

In the face of contradiction, one must make distinctions. The Panaetian distinction between the four personae and realms of duty allows us to mitigate the tension in E40. In *On Duties (De Officii)* Cicero reports Panaetius' view as having duties in light of (i) our being rational creatures (universal duties), (ii)

our having special individual endowments, (iii) our having circumstances of chance provide social responsibilities, and (iv) our choices and volitions (*De Offic*. I.30.107-115). Epictetus inherits and endorses this four-personae view in his heuristic for discovering duties in D.2.10. First, one is a rational being. Next, one is a rational being with unique capacities. Further, one is a rational being with unique capacities with unique familial relations and specific citizenship. Finally one is a rational being with unique capacities, relations, and with a history of having made specific choices. Once one has completed this heuristic, one can see one's duties more clearly, since: "Each of these designators, when duly considered, always suggests the acts appropriate to it" (D 2.10.12). Epictetus, then, runs the reasoning for a man as recognizing that he is rational, has certain capacities for speaking, is a brother, son and a Roman, and so has responsibilities as having taken on the role as councilman, husband, father, and friend. Once he considers these roles, he sees his duties.

The same, it stands to reason, Epictetus would say goes for a woman. And so, for some woman, we might say she sees herself as a rational being, with a family, a city, chosen friends, and household responsibilities. So a woman's duties are determined by her relationships, as Epictetus makes clear in E30. For the woman, her relationships are, given her social role and her opportunities, overwhelmingly familial.

E40, then, stands as a corrective for young women who are confused about what their true roles are. In the same way that many a young man may be distracted from his duties by interest in the esteem of others, or wealth or pleasure, so, too, may a young woman be distracted by sexual interest. The apparent contradiction, then, can be resolved with the Stoic notion of indexing duties to given social roles, and as a consequence, one may square the program of consciousness-raising and progressive educational opportunities for women with the conservative views about women's duties and social opportunities. Educational reform and consciousness raising is about making sure that women can live up to their responsibilities as rational creatures, but that has no bearing on what social opportunities they are to be offered. Given the Four-Personae view, these are separate spheres of the person. In fact, Stoic value theory consistently maintains the distinction between social standing and opportunities and the goods of the soul. The goods of the soul are what matter, and the rest are incidental to true virtue and happiness. Once one has the virtuous soul, one plays the part into which one is cast – no matter if it is a Caesar or a slave, a merchant or a beggar, a philosopher or a housewife. The question, now, is whether this consistency is bought at the price of moral soundness.

III

There are at least four aspects of Stoic theory which can be made foci for concern for feminists. We will discuss these four aspects in order of ascending importance. The first thing to note is that the Stoics in general, and Epictetus in

particular, are addressing only men – we call this the *limited audience* problem. In E40, Epictetus advises men to "fix [their] attention" to make women understand that they are valued for their modesty (E40). Musonius Rufus addresses fathers and husbands regarding the education of daughters and wives, but never talks to the daughters and wives themselves (Nussbaum 2002, 303). The claim, by those who wish to defend the Stoics as feminists, is that the Stoics allow both men and women to philosophize. Women share in the same virtue as men and the study of philosophy allows them to realize this virtue (Hill 2001, 19).[4] It remains a mystery, however, why the principal Stoic works are addressed exclusively to men if women are equally able to participate in philosophy.[5]

One could argue that the male audience is explained by precedent alone, although we're not convinced – especially considering some of the Stoic's blatantly sexist sentiments. Cicero and Seneca, for example, use feminine adjectives (like *muliebris*) to denote moral failings and masculine adjectives (like *virilis*) to denote praiseworthy actions (Manning 1973, 171). Seneca goes so far as to say that among those who do philosophy, Stoics are the only ones who consistently argue and think *like men* (*Ad Serenum* II.1). Musonius Rufus argues that men are superior to women and natural rulers, while women are naturally ruled (as noted by Engel 2003, 281; Nussbaum 2002, 303). Musonius also argues that a woman should "learn to love her children more than her life," and Epictetus dismisses Epicureanism as not befitting even women (Stob. *Anthol*. III 6,57). In short, the fact that Stoics claim that women share the same virtue as men is not enough to convince us of their egalitarianism. Women are never addressed – all advice about women is addressed to men – and when women are cited at all it is often not in a flattering way. Our argument here, however, is largely one about what to infer about the Stoics, *de facto*, from silences in their writing. More work is needed to make our argument stick.

If we return to Epictetus in E40, we may be tempted to credit him for pointing out the impact of oppression on the oppressed. He tells us that women internalize their sexual objectification and construct their self-image, and self-

[4] Note here that Hill does not acknowledge the point we accept, that equal access to virtue in one sphere of life does not make one equal in other spheres. Equal access to philosophical education (even if we grant that this is something endorsed by the Stoics) does not ensure, most importantly, social equality.

[5] The notable exception is Seneca. Two of his letters are addressed to women: *Ad Marciam* is a letter of consolation to Empress Livia's friend and daughter of the Republican historican, Cremetius Cordus. *Ad Helviam* is a letter Seneca addressed to his own mother while he lived in exile. In these, Seneca is clear that he regards women as natural equals, but also those with unique occasions for vice – he even opens his letter to Marcia acknowledging that she is an exception to the wide majority of women who suffer from the "feminine weakness of mind (infirmitate muliebris animi)" (*Ad Marc*. VI.1) That Seneca must acknowledge the fact that he is addressing a woman in philosophical manner is evidence that it is the exception that proves the rule.

worth, accordingly. We might see this passage as consistent with a larger Stoic argument regarding the common humanity of men and women alike. But note that in this passage Epictetus is also criticizing the women who construct their self-image based on their own objectification. We might call this a sort of *blaming the victim* problem. It is the women, after all, who upon realizing their only hope is to persuade men to "go to bed" with them, begin to "make themselves up and place all their hopes on that" (E40). They are participants in their own oppression. It is also important to note, in relation to our first worry, that Epictetus is urging men to inform women of their true worth so that women will no longer be willing participants in their own objectification. In this sense, women's salvation is still dependent on men.

This brings us to our third worry – the worry about *continued subordination*. In Epictetus and many other Stoics (including Cicero, Musonius, and Seneca) women are viewed as ultimately dependent on men for success. Although Epictetus does raise women above the status of sex-object to one of modesty (an improvement to be sure) he still defines women's goodness in terms of how they appear to men. Note that he doesn't argue that women are valued for their modesty, but for "*displaying* (*phainesthai)* decent, modest and discreet behavior" (E40). Lest we think this is a merely linguistic quibble – consider Musonius Rufus, who argues that educating women will turn out to be good for the husband (Stob. *Anthol*. 4.28.20, and see Nussbaum 2002, 303); or Hierocles, who argues that women should know how to labor efficiently so that they might "fulfill the orders of the master of the house" (Stob. *Anthol*. 4.28.21, and see Engel 2003, 284). Educating women is thus a good thing primarily for men. Displaying modest behavior is a trait that men value in women. But it's important to point out here, that even if the argument were made that modesty is good for women themselves, the Stoics would simply be providing another role to force upon females.

So far we have considered three worries – the limited audience problem, the blaming the victim problem, and the continued subordination problem – which lead us to question the Stoics' commitment to feminism. All three of these worries are united by our fourth, and most significant, concern – the *social status* problem. Our argument for the consistency of the two aspects of Stoic doctrine (the program of consciousness-raising and progressive educational opportunities, and the conservative views about women's social roles) relied on the notion of separate spheres. Stoic value theory maintains the distinction between social standing and the goods of the soul; so although women share the same virtue as men, and thus share this good of the soul, they must also exercise this virtue in a completely different social arena than men. While men are permitted to do politics, for example, women must stay in the home.

What the Stoics fail to completely acknowledge, however, is the way that goods of the soul rely on and are not completely separable from, social standing. Nussbaum argues of the Stoics, for example, that they fail to "understand the

extent to which human dignity and self-respect require support from the social world" (Nussbaum 2002, 302). The aspiration of Stoic ethics is self-sufficiency (*autarkeia*), but the Stoics consistently acknowledged that this aspiration is regularly just that – aspirational. Seneca notes that having enough rest helps with controlling anger (*De Tranq.* XVII.5; *De Ira* III.ix.5), that children should not be subjected to degrading treatment or made to be excessively servile or submissive (*De Ira* II.xxi.4). Seneca even goes so far as to counsel that it is wise to avoid people and conditions that will provoke irritation (*De Ira* III.vi.3 & III.viii.3). And finally, Seneca holds that station is preferable for and conducive of virtue, since the good soul has free play to express itself as a judge rather than as the judged, the benefactor instead of the beggar (*De Clem.* I.v.3). The reality is that though the Stoic goal is self-sufficiency, we are not independent creatures, and the contingencies of our lives have immense consequences for our opportunities for virtue. It is, then, no coincidence that Marcus' opening book of the *Meditations* is a long list of people who had been good teachers and exemplars. The implication is that without them, he would not have had such virtue.

The lived reality of women's lives, for the Stoics, is that they are good and efficient homemakers. An education in virtue simply allows them to fulfill this purpose better. And while many contemporary women might embrace homemaking as a worthwhile lifestyle, they enjoy something completely lacking from the Stoic story – *choice*. The very same separation of spheres that allows the Stoics to be consistent thus condemns them on the feminist question. Consistently hearing that one is not capable of choosing well or having any opportunity to do so destroys whatever capacity one has to exercise one's capacities. The Stoics regularly recognized that exercises were necessary for the perfection of judgment, desire, and action – but if one is never offered those opportunities to ever exercise them, one cannot fully develop one's capacities for rationality and virtue. That should be troubling.

Why is it, then, that so many contemporary scholars continue to consider the Stoics feminists or quasi-feminists? For a possible solution, we suggest looking at the definition of feminism at play. Lisa Hill, for example, defines feminism as "A view of women as a distinct sociological group for which there are established patterns of behavior, special legal and legislative restrictions, and customarily defined roles..." (Hill 2001, 14). This definition says nothing about choice or social status, and so Hill can conclude that Stoicism is consistent with feminism. We argue, on the contrary, that any acceptable definition of feminism must include an aspect of choice for women, and must say something about reforming the isolation of women to the private sphere. When this definition of feminism is used, it is clear to see that the Stoics are not, in fact, feminists. And on the assumption that the liberalism requirement is right, then any Stoic view that runs afoul of the choice principle is normatively unsound.

IV

By analogy, consider the question of Plato's *Republic* as a politically feminist document. The Principle of Specialization running the beautiful city is that peoples' jobs are indexed to abilities, and women with the same abilities as men should play the same social roles as men (456a). Both men and women can use reason to see the Forms, so both should have the opportunity to pursue that educational opportunity. As a consequence, women can be full guardians (*phulakes pantiles*). Glaucon, horrified at the prospect of women leading (and the sight of old women naked in the Palestra), objects, but Socrates reminds him that it is a matter of justice that jobs and political opportunity be available to those with ability, regardless of sex. So the beautiful city, as a matter of justice, has equal opportunities for women. This is not only a utopia, but, it seems, a *feminist* utopia (See, for example, Vlastos 1997).

Despite these initial progressive sentiments, Plato, like the Stoics, has an uneven track record with women, too. First and foremost is the fact that, despite claiming that men and women are equal in some capacities, they are not so equal in others. Socrates and his interlocutors agree men have greater mental and physical prowess (455b), and so men perform most jobs better than women (455d), even those widely considered to be womens' work. Plato, additionally, is no stranger to the offhand misogynist crack. So vices, like that of being exceptionally emotional, are expected in women (387e), and even if women are able to do philosophy, we don't expect them to be consistently good at it (455b).

Independent of Plato's rhetoric is the final point that, as Julia Annas notes, there is "no reference to women's desires or needs" (1976, 311). Women (and men, alike) are given social roles that are determined by their natures, and then expected to perform them without question. Women's roles are made equal, not for their benefit, but for the state's. The case against Platonic feminism is to be made along the lines of the question: *what kind of role does an individual's choice make in determining that person's life*?

If the answer is that individual choice plays none for women (or men, for that matter), then our view is that this makes the view non-feminist. We will defend this view, what we call *the liberal core of feminism*, later. But for now, notice that the assessment of Plato as feminist depends on the status of *individual rights* as a prerequisite for feminism. If one takes, one might say, equality of treatment and opportunity regardless of sex as the only requirement for feminism, then Plato counts. If one requires the further liberal rider of individual choices being respected, then Plato fails.

It is an irony of intellectual history that Epictetus, too, sees the matter with Platonic feminism clearly. Stobaeus reports that Epictetus criticized the women of Rome for carrying around copies of Plato's *Republic*, thinking that the Platonic political vision would be something that would liberate them and create a "community for women" (*koinas… tas gounaikas*). Instead of proposing that they be liberated from their encumbering marriages to particular men, Epictetus

holds that Plato "first abolishes that kind of marriage, and introduces another kind, to the state, in its place" (Stob. *Anthol*.6.58 – Epictetus Fragment 15).

Plato is the ancient touchstone for our reconstruction. Plato is a proto-feminist, but one who failed to think through the demands of feminism as a justice movement in the right way. The Stoics, as we've argued, fail in similar fashion. Ancient ethical and political thought has a long line of failure in thinking about individual rights, and the figures around whom we reconstruct Stoic ethics and politics recapitulate the oversight. But Stoicism as a broader frame of thought, we might say as a philosophical *Zeitgeist*, has the capacity to construct an approximation of liberal feminist commitments. What follows is a brief overview of what tools are available within the Stoic tradition to frame the thought, and then we will close with a defense of the liberal requirement we've deployed for genuine political feminism.

V

The Stoic tradition was not merely a set of academic philosophical doctrines. It was, particularly by the time of the Imperial Stoa, a cognitive paradigm, and one that was the default for intellectuals (Cf. Shaw 1985). Testament to this fact is the phenomenon of contrast by all those who were non-Stoic philosophers in the period. The most important job for a neo-Platonist, Epicurean, or Pyrrhonist is to make it clear where the view on offer critically differs from or overlaps with the prevailing rough set of Stoic views.[6] The reality is that Stoicism had its *dogmata*; however, they were a rough but familiar list. Cicero's digest in *Paradoxa Stoicorum* is representative, but it seems clear that Marcus, for example, diverges widely on the singularity of vices (M 2.10), and the reality is that by the time of Seneca's *De Clementia*, Stoic politics had come a long way from the Cynicizing idealistic fervor of Zeno of Citium's *Republic.* Stoicism was, it seems, more a rough and ready range of intellectual options all bearing family resemblances, but nevertheless allowing for wide disagreement.

It is within this range of intellectual diversity for the Stoic tradition that we pose *Stoic feminism*. The place to start is with the Stoic notion of natural law. Plutarch describes Zeno's *Republic* as based on a notion of a "law common to all (*ho nomos ho koinos*)" (Plutarch *De Alexandri Magni Fortuna aut Virtute:* 392 a-b). This notion of natural law was expressed in Stoic philosophy in projects ranging from logic and the theory of inference being universal[7] to the political aspirations of cosmopolitanism.[8] The universe and we, within our minds and as

[6] See Sextus' keenness on making it clear how his criticisms of dogmatism first and foremost bear on the Stoics (e.g., PH III.181). See Simplicius' commentary on the *Enchiridion* for the neoPlatonist urgency of appropriating Stoic doctrines (3,5).

[7] See, for example, the story related by Sextus Empiricus of Chrysippus attributing the use of disjunctive syllogism to his dog (PH I.69).

[8] As we see most clearly expressed in Marcus Aurelius, who claims that he is a citizen both of Rome and the world.

bodies, are governed by *the logos*. The objective of Stoic life, then, is to come to understand and live in accord with this *logos*. We thereby, come to make ourselves *at home* in this world. Diogenes Laertius calls this *oikeiosis*, the process of coming to live in accord with our nature (literally, home-making) (DL VII.85).

This core concept, that of living in accord with nature, has a significant ambiguity. On the one hand, *accordance with nature* can simply mean being in accord with *what is*. One, as Epictetus says, wills in accord with how one finds the world (E8). Call this *thin naturalism.* On the other hand, nature has a teleological, normative element to it. Injustice, for example, is unnatural (Marcus *Meditations* 2.16 & 9.1). Call this *teleological naturalism.*[9]

It is in this duality of *accordance with nature* that we see, first, the tools for diagnosing why so many Stoics failed the critical program for feminism, and second, that there are tools for developing the progressive feminist program. In short, if *oikeiosis* requires living in accord with what is, the thin naturalism, then programs of drastic cultural change or criticism are objectionable. However, if we see natural teleology in the divine reason in each human, then there is reason for cultural criticism. Cultural norms that contravene or degrade the natural dignity of rational human choice (*prohairesis*) deserve criticism and should be changed.

The Stoics' uneven track record with women is recapitulated by their treatment of slavery. On the one hand, the Stoics widely decried the treatment of slaves and even the very institution of subjecting another human to domination. Epictetus seems to even propose that there are some activities (e.g., holding a chamber pot) that are beneath human dignity and should be refused even at the price of a lethal beating (D I.2). On the other hand, the Stoics never moved beyond this mere theoretical criticism. For sure, Marcus Aurelius made manumission easier, but he took no steps toward restricting slavery. And Epictetus, as reported by Simplicius, despite having been a slave himself, takes a slave to nurse the child of a neighbor who was about to expose it (116,50 – Brennan 2002, 95).

But the uneven track record on slavery for the Stoics needn't undercut their progressive line of thought. The conservative thinly naturalistic viewpoint is relevant only for framing feasible policy change, but not the direction in which policy should be changed.

The same, we hold, is the case for feminism. The stoic natural-teleological view is that women have rational natures and a capacity for reasoned choice. The consequence is that from the perspective of the goods relevant to moral goodness, women are men's equals and deserve the same respect and dignity that men are afforded. And this is precisely why Musonius Rufus holds that women deserve to be taught philosophy, why Seneca holds that women have the same capacities for virtue as men, and why Epictetus criticizes the sexualization

[9] See Susan Ford Wiltshire (1992) for a discussion of how this duality of the notion of naturalness yielded a robust notion of natural law and the notion of natural rights.

of young women. What is valuable in women, their capacity for rational choice, is not being respected. Culture criticism is necessary in those cases, and the Stoics consistently came to criticize their own cultures for these failings. The problem was always what they proposed in its place.

But what of Stoic endurance, what of the requirements that the Stoic distinguish between internals and externals and be resigned to what is? Epictetus proposes that we should not look to have events happen as we want them to, but to want them to happen as they do (E8). Stoic ethics, as David Engel notes, has a core commitment to resignation, and so, as Engel takes it, to conservatism:

> [T]he Stoics' emphasis on the moral indifference of one's external conditions makes it unlikely that they ever advocated a more prominent role for women in political live. *It probably also makes feminism and Stoicism not just contingently, but essentially incompatible* (2003, 288, emphasis added)

But this, we hold, is a non-sequitur. Engel's view has the Stoic feminist be indifferent to externals in the sense that they are not reasons for action. However, the Stoic, again, identifies *duties* and then acts in ways pursuant of them. Insofar as our duties to each other as fellow rational creatures requires respecting each other's choices (that which is the expression of our moral purposes – our *prohairesis*), then we do have responsibilities to each other. Just because externals are indifferent to us does not mean that we have no moral reason to nevertheless act in ways pursuant of external justice.

A further line of argument is necessary to properly rebut Engel's point. Stoic value theory runs that social standing cannot affect one's virtue, and since virtue is sufficient for the good life, unequal treatment is not really harm. Again, we've argued that *there is a difference between recognizing that one is not harmed in being treated unjustly and being complicit with unequal treatment or participating in it*. Surely it is a harm to be knowingly unjust, but harm to the virtue of the person complicit with or participating in the injustice. And so, Stoic value theory requires an active life of advocating for justice, equal opportunity and respecting the choices made by others. Our strategy, then, is to say that the doctrine of the moral indifference of externals is orthogonal to the requirements of Stoic feminism.

This said, the doctrine of moral indifference to externals can still be useful to, instead of contrary to the purpose of, the Stoic feminist. Consider: the Stoic can have a critique of the institution of slavery or any other unjust treatment of people, but then also have strategies for life that makes it so that when injustices happen to us, we can endure them. Epictetus prepares to go to the baths by readying himself for the rude and raucous behavior of others (E4). When he goes and is splashed or has someone act inappropriately around him, he must understand that he signed up for the whole experience. And so he is ready to endure what must be endured. But this is not an *endorsement* of the rude or raucous behavior. No Stoic would endorse such actions, but would criticize them.

The same should go for the Stoic feminist – we identify the correct conditions for justice, but we prepare ourselves for when injustice arrives. There is, then, living in accord with what is (thin naturalism's acceptance of what is), and living in accord with what natural reason requires (recognizing the ways one's culture can fail to manifest divine reason).

VI

So far we have argued that the Stoics' uneven track record can be resolved, and that the two trends within Stoicism (the progressive view about women's capacities and the conservative view about women's roles) are consistent. We have also argued that Stoicism's two trends *can be* aligned in a way which is morally sound. Our argument here focused on what we call *the liberal core of feminism*. In this final section, we offer a brief defense of this idea.

We argue that respect for individual choice is a *sine qua non* for any feminist theory, and it is this respect which constitutes feminism's liberal core. The respect for individual choice is entailed by the moral equality of persons, that is, liberalism's stipulation that people are free and equal, and capable of the rational choice of ends and conceptions of the good. In other words, each person should be viewed as a "self-authenticating source of valid claims" (Rawls 2005, 32). Because this respect for individual choice is entailed by the commitment to moral equality, any theory grounded on the moral equality of persons must respect individual choice. Feminism, which is grounded on the moral equality of *women*, must therefore similarly respect women's choices. Not respecting women's choices denies their status as free and equal, and given that this result is unacceptable for any feminist account, we argue that feminism should readily accept the liberal rider to respect individual choices.

To put the point another way, liberalism is committed to the idea that all people should enjoy both personal and political autonomy; that is, each person should be able to choose the kind of life she wants to lead, and she should similarly be able to (at least partially) determine the circumstances under which she leads this life. Feminists, too, are committed to promoting autonomy, more specifically, the autonomy of women. Historically, feminists have been concerned to free women from the forces of misogyny and patriarchy – forces which denied women the power of choice over their own lives. It seems, then, that feminism already has (at least minimally) a liberal core.

But our argument here rests on the stronger claim that feminist theories must necessarily have this liberal core. In order for a theory to be considered feminist, it *must* respect individual choices. Why? Because to do otherwise is to fail to recognize that women have rational natures and the capacity for reasoned choice. To establish a set of norms for women which dictates what sorts of roles or actions are appropriate for them *as women*, is to deny their equal moral status and their status as rational choosers. This result should be unacceptable to any feminist. Insofar, then, as feminism is committed to women's freedom from the

dominating forces of patriarchy, insofar as it is committed to the idea that women share men's capacity to choose for themselves, it must be committed to the respect of individual choice.

Under this conception of feminism, the Stoicism of the individual Stoics on offer fails. Recall that, although Epictetus criticizes the cultural norm that views women as sex-objects, he proposes another norm in its place. This norm – of appearing modest – is appropriate for women based on their nature *as women*; individual choice plays no role in determining their lives. But Stoicism as a philosophical *Zeitgeist* does, in fact, have the capacity to be a feminist theory. This capacity, as we've shown, stems from Stoicism's respect for rational human choice (*prohairesis*) – the same respect that constitutes the liberal core of feminism.

VII

Our conclusion is, then, threefold. First, that feminism requires respecting individual choice. We have called this *the liberal core of feminism.* Second, that *the Stoics*, despite their feminist inclinations (or we might say, *protofeminism*), failed to respect the autonomy of individual women. This, we've argued, is not inconsistent with their Stoicism, but is morally unacceptable. This is what we've called *Stoicism's Uneven Track Record.* Third, and finally, we have argued that despite the fact that the individual Stoics themselves failed the liberal requirement, Stoicism as a philosophical program is not inherently anti-liberal (and thereby anti-feminist). We've provided a sketch of what this liberal Stoicism looks like. A Stoicism 2.0, if you like. The liberal Stoicism we've proposed respects autonomy, but it recognizes the fact that the world is not ideal, and so there must be the familiar Stoic virtues of endurance. And these virtues of endurance needn't be inherently socially conservative or misogynist.

References

Annas, Julia. 1976. "Plato's Republic and Feminism." *Philosophy* 51: 307-321.

Arnold, Edward V. 1911. *Roman Stoicism.* Books for Libraries Press: New York.

Asmis, Elizabeth. 1996. "The Stoics on Women." In *Feminism in Ancient Philosophy,* ed. Julie K. Ward, 68-92. Routledge: New York.

Bates, Carrie. 2014. "Self-Identity in Epictetus: Rationality and Role." In *Epictetus: His Continuing Influence and Contemporary Relevance*, ed. D.R. Gordon and D.B. Suits, 151-163. Rochester: RIT Press.

Cicero. 1942. *Paradoxa Stoicorum,* trans. H. Rackham. Cambridge:
Harvard University Press.

Cicero. 2005. *De Officiis*, trans. Walter Miller. Cambridge: Harvard University Press.

Diogenes Laertius. 2000. *Lives of Eminent Philosophers*, trans. R.D. Hicks. Cambridge: Harvard University Press.

Engel, David M. 2003. "Women's Role in the Home and the State: Stoic Theory Reconsidered." *Harvard Studies in Classical Philology* 101: 267-288.

Epictetus. 1998. *Discourses,* trans. W.A. Oldfather. Cambridge: Harvard University Press.

Epictetus. 1983. *The Enchiridion,* trans. Nicholas P. White. Indianapolis: Hackett.

Hill, Lisa. 2001. "The First Wave of Feminism: Were the Stoics Feminists?" *History of Political Thought* 22: 13-40.

Manning, C.E. 1973. "Seneca and the Stoics on the Equality of the Sexes." *Mnemosyne* 26: 170-7.

Marcus Aurelius. 2003. *Meditations,* trans. Gregory Hays. New York: Modern Library.

Musonius Rufus. 2011. *Lectures and Sayings,* trans. Cynthia King. CreateSpace.com.

Nussbaum, Martha. 2002. "The Incomplete Feminism of Musonius Rufus." In *The Sleep of Reason*, ed. M. C. Nussbaum and J. SHivola, 283-325. University of Chicago Press: Chicago.

Pappas, Nickolas. 1995. *Plato and the* Republic. New York: Routledge.

Plato. 1937. *Republic,* trans. Paul Shorey. Cambridge: Harvard University Press.

Pomeroy, Sarah B. 1975. *Goddesses, Whores, Wives and Slaves.* Shocken Book: New York.

Rawls, John. 2005. *Political Liberalism*. New York: Columbia University Press.

Seneca. 1928. *Moral Essays,* Vol 1, trans. John Basore. Cambridge: Harvard University Press.

Seneca. 2006. *Moral Essays,* Vol 2, trans. John Basore. Cambridge: Harvard University Press.

Shaw, B.D. 1985. "The Divine Economy: Stoicism as Ideology." *Latomus* 44: 16-54.

Vlastos, Gregory. 1997. "Was Plato a Feminist?" In *Plato's* Republic: *Critical Essays, e*d. R. Kraut, 115-128. Rowman and Littlefield: New York.

Witlshire, Susan Ford. 1992. *Greece, Rome and the Bill of Rights.* Norman: University of Oklahoma Press.

Søren Kierkegaard's *Repetition*. Existence in Motion[1]

Ionuţ-Alexandru Bârliba

Abstract: This article tries to make sense of the concept of repetition in Søren Kierkegaard's works. According to Kierkegaard repetition is a temporal movement of existence. What is repetition and what is its meaning for human existence? In answering this question the Danish philosopher depicts repetition by comparing three different approaches to life. Throughout the article I try to develop a coherent argument on 'the new philosophical category' by analysing the three types of repetition and their corresponding human prototypes. I consider repetition a key concept in summarizing Kierkegaard's theory of existence, where existence pictures the becoming of the human-self that follows several stages. Constantin Constantius's repetition is an unsuccessful attempt, an aesthetic expression of human-life. The young lover's repetition is spiritual, albeit not yet authentic, religious, but more poetic, even if he regains his self. Only Job's repetition is an authentic movement of existence, an expression of a spiritual trial and of genuine faith.

Keywords: Søren Kierkegaard, repetition, recollection, self, existence, motion

In 1843, Søren Kierkegaard published *Repetition*, under the pseudonym Constantin Constantius, 'an insane book' (as was designated by the author of *The Concept of Dread*, Virgilius Haufniensis), followed by, in the very same day, the *Fear and Trembling* volume, signed by Johannes de Silentio. After just a few weeks, Kierkegaard published a piece on *The Book of Job* in his *Edifying Discourses* series. The coherence between the three works is hardly fortuitous, a feature emphasized by their topics.

Thus, in *Fear and Trembling*, Abraham, 'the father of faith,' must sacrifice his only son only to get him back later. Between the pages of *Repetition*, we also learn, on the one hand, about Constantin Constantius' voyage in Berlin, undertook with the hope of retrieving some of the contentment experienced during a previous visit, and on the other hand, we meet the unfortunate case of a young lover who ended the engagement with his fiancée, only to languish over her afterwards. Also, *Repetition* concisely brings up the sufferings of a biblical character, Job.

[1] This paper is 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/159/1.5/133675.

As one can easily notice, all the above-mentioned characters are connected by the same underlying tendency, the need to re-experience a previous condition. The degree to which such a reliving (a re-turn) is possible and most of all, finding out whether such a temporal motion is possible, can be achieved only through the 'new philosophical category' framed by Kierkegaard through Constantin Constantius' voice, namely through repetition.

Repetition is one of the most problematic terms found in Kierkegaard's works. At the beginning of the previous century, Walter Lowrie confessed that one cannot find a more important and, in the same time, more confusing term in Kierkegaard's writings, as the term repetition (Lowrie 1938, 630). Among other reasons, this confusion is also generated by the merely rigorous manner in which Kierkegaard chooses to approach the discussion on repetition.

As a genre, *Repetition* can be rather described as a love story or a psychological novel; concurrently, the book can be regarded as a diary with inserted letters.[2] Anyhow, Repetition is not an objective philosophical treatise; its reading does not necessarily lead to the elucidation of the concept proposed for inquiry. Nevertheless, one cannot reach the conclusion that *Repetition* is completely deprived of a conceptual framework that could lead to the conclusion that the meaning of the concept remains completely undisclosed.

To all intents and purposes, one could find two different routes for approaching repetition, two modes that will also organize the contents of the following pages. In other words, I identified two alternative routes: a conceptual approach and an experimental one. Examining concepts with already established philosophical grounds, such as the concept of recollection, motion or mediation, connected by Kierkegaard to the idea of repetition, and to the analysis of the existential examples offered by the characters from *Repetition*, should ultimately throw light upon this new philosophical category proposed by Søren Kierkegaard.

Paraphrasing Constantin Constantius, in the following pages, I will try to discover what repetition is, whether it is possible and what its (existential) meaning is. By clarifying these aspects, I will be able to reach some interesting conclusions about the role of repetition in the process of self-becoming, in the light of the three stages of individual existence, proposed by Kierkegaard. Concurrently, I will be able to highlight Kierkegaard's particular approach of the concept of time.

1. Recollection vs Repetition

Constantin Constantius introduces the 'the new category that will be discovered' (Kierkegaard 1983, 148), as the author described the concept of repetition, during a rather confusing discussion about motion and recollection in Ancient Greeks:

[2] Besides, the subtitle of *Repetition* is *A venture in experimental psychology.*

> When the Eleatics denied motion, Diogenes, as everyone knows, came forward as an opponent. He literally did come forward, because he did not say a word but merely paced back and forth a few times, thereby assuming that he had sufficiently refuted them. When I was occupied for some time, at least on occasion, with the question of repetition – whether or not it is possible, what importance it has, whether something gains or losses in being repeated – I suddenly had the thought: You can, after all, take a trip to Berlin; you have been there before, and now you can prove to yourself whether a repetition is possible and what importance it has (Kierkegaard 1983, 131).

After evoking the hilariously ended dispute between Diogenes and the Eleatics, Constantin Constantius confessed that he has been occasionally (thus, rather detached) concerned with this problem, a problem that he preferred not to call motion, but repetition. The natural transition drawn by Constantin from one passage to another leaves us with the impression that the anecdote about Diogenes and the Eleatics and his preoccupation with repetition are concerned with the one and the same thing; a not at all wrong impression, further confirmed by the introduction in the text of the second invoked mention, recollection. "Repetition and recollection are the same motion, except in opposite directions" (Kierkegaard 1983, 131). Could Constantin Constantius have tried an experiment similar to the one employed by Diogenes to convince himself of the possibility of the motion? Hard to say, although it is true that, to test the possibility of repetition, he took a trip to Berlin, a few months after his first, very enjoyable visit to the same city.

If repetition is motion, one has to decide its direction. A useful clue is identifiable hereinafter, following the fragment cited above:

> Repetition and recollection are the same motion, except in opposite directions, for what is recollected has been, is repeated backward, whereas genuine repetition is recollected forward (Kierkegaard 1983, 131).

Because both are movements, a first step in uncovering the concept of repetition is the act of 'remembering' the meaning of recollection in Greeks. The doctrine of remembrance, introduced by Socrates in the Platonist dialogue *Menon*, has its origin in the following controversial topic:

> A man cannot enquire either about that which he knows, or about that which he does not know; for if he knows, he has no need to enquire; and if not, he cannot; for he does not know the very subject about he is to enquire (Jowett 1892, 40).

The reasoning thus phrased seems to inhibit any possibility of knowledge or, at least, any learning attempt. Socrates uncovers this paradox, by shifting the focus from the object of knowing on the knowledge itself, more specifically, on understanding the ways in which knowing manifests in the person. Thus, through a series of sequential interrogations, Socrates manages to bring to light a slave's knowledge in Geometry; the kind of knowledge that the slave was never acquainted with and which was thus possessed by his soul since from the beginning. However, they could be brought to light through Socrates' mediating

intervention. Nevertheless, Socrates does not teach Geometry to the slave. He only brings the decisive occasion, thanks to which, Menon's slave reveals his acquired knowledge, either in a forthcoming existence, or knowledge that he had since the beginning, latently manifested in his soul. In this sense, Socrates' conclusion is "… for all enquiry and all learning is but recollection" (Jowett 1892, 40).

What needs to be underlined here, of relevance for the present discussion, is the fact that recollection is a form of knowledge, through which we generally try to find out that "the truth of all things always existed in the soul" (Jowett 1892, 47).

If recollection is a way of knowledge that recuperated a truth that already exists, repetition should recapture a forthcoming truth. What becomes increasingly clear is the fact that the movement proposed by the two forms of knowledge is not one that happens along a 'spatial' axis as it is, in a first denotation, indicated by Diogenes' solution), but one that follows a temporal axis.[3] Remembrance points to a past moment. Anyone who can reach an earlier reality through recollection travels back into the past. In a Platonist acceptance, remembrance indicates a static previous time, unidentifiable and eternal. Repetition, in turn, as a movement opposed to remembering, is directed towards a future time.[4]

However, the act of repetition requires the existence of a 'repeatable,' re-actualizing reality; otherwise, repetition would have the meaning of a simple feeling of hope or a planning act. Trying to acknowledge, at this first level, the meaning of repetition in Kierkegaard's vision, I could state that: repetition is that human need of bringing into actuality, of reviving, the joy produced by the presence of an object (broadly defined) or an endeared lost person, at a certain

[3] For instance, in the supplement to the Hong & Hong edition of the *Repetition*, we can read the following phrase:
Movement is dialectical, not only with respect to space (in which sense it occupied Heraclitus and the Eleatics and later was so much used and misused by the Sceptics), but also with respect to time (Kierkegaard 1983, 309).

[4] The distinction operated by Kierkegaard between the way recollection and repetition act can be related to the author's distinction between the two types of religiosity (A and B), that define the last stage of his existential dialectics, also corresponding to the movement of the eternal resignation and the movement of belief by the virtue of the absurd, concepts that are central for the volume *Fear and Trembling*. It is important to note that, similarly to the infinite resignation, recollection is an action whose success resides in the hands of the person, while the success of repetition requires the involvement of the divine, who will find not only the possibility of a repetition but also the reaction of the person, through faith; because, as we will find out, repetition is also a religious movement by virtue of the absurd. Also, the two types of religiosity present two distinct ways of accessing the eternal, and they could be identified as recollection, through which the eternal is regained and repetition, through which the eternal is projected (believing by virtue of the absurd). Anyhow, Kierkegaard identifies religiosity A as rather specific to the pagan, greek world, while repetition is a religious movement by virtue of the absurd, thus identifiable with religiosity B.

moment. What brings specificity to this new category in Kierkegaard's thought is the paradoxical belief[5] in regaining what was lost, despite the competing evidence.[6] Repetition brings into present a past reality, a reality of whose truth we will be edified only in an indeterminate, eternal future.[7]

At this point, it is important that we direct our attention once again towards the anecdote of Diogenes. Commentators such as Claire Carlisle point to the fact that the fragment reveals the opposition between ideas (philosophy) and movement, opposition that, in Kierkegaard's terms, parallels the antinomy between recollection and repetition. Thus, if remembrance is a form of knowledge that regains an already existing truth, as *Idea*, repetition is a movement of becoming, of truth's coming into existence (Carlisle 2005a, 522). For that matter, Kierkegaard himself operates this distinction, through Constantin Constantius voice:

> When the Greeks said that all knowing is recollecting, they said that all existence, which is, has been; when one says that life is a repetition, one says: actually, which has been, now comes into existence (Kierkegaard 183, 149).

Recollection is a form of knowledge through which the person actualizes or brings into existence a permanent truth. Repetition is, however, an act in motion, by which truth becomes known, together with and through the repetition experience. The event which is 'to be repeated' is not a static ideality, fixed in a past eternity, which imprints a model of reality. The object that 'is to be repeated' represents only the starting point or motivation for becoming. If recollection regards knowing the (already existing) truth, repetition requires its discovery; in other words, its reception through belief.

A movement (recollection) is static, it does not bring anything new, and it does not alter the individual's existence. The alternative movement (repetition) transforms the individual's existence; it re-actualizes it by attaching a meaning and coherence to it.

---

[5] Søren Kierkegaard conceives paradox as the circumstance in which logic and reason are obsolete, defied, characteristics that confirm its validity. In addition to this idea, Kierkegaard also notes:
When repetition is defined in that way, it is: transcendent, a religious movement by virtue of the absurd- when the borderline of the wondrous is reached, eternity is the true repetition (Kierkegaard, 1983, 305).

[6] As I will show in the following pages, there are several kinds of repetition, some of them failing not only as imposed motivation but due to their execution; others, such as those enacted by Job or Abraham, are successful precisely due to the mentioned elements.

[7] At this point, Clare Carlisle's observation is also noteworthy:
So, both recollection and repetition are movements of truth: the former moves towards a past eternity, and the latter moves towards a future eternity (Carlisle 2005a, 525-526).

2. Repetition is a 'Transcendental Movement'

What has been defined as repetition up to this point is, according to Kierkegaard, inadequately interpreted as mediation (ger. *Vermittlung*) by the Hegelian thought. In one of the strictly philosophical fragments in *Repetition*, Constantin asserts "that repetition proper is what has mistakenly been called mediation" (Kierkegaard 1983, 148). In order to better understand Kierkegaard's view on mediation, one needs to focus on *The Concept of Dread*, a book that is generally recognized as the conceptual framework for the colloquial volume signed by Constantin Constantius.

The central point in Kierkegaard's critique of the Hegelian philosophy in this context is the claim about the possibility of motion in logic. Kierkegaard's argumentation starts from the observation that (Hegelian) logic cannot regard reality as its ultimate concept. Logic cannot find its culminating point in reality, in the real existence.

What makes this relation impossible is the category of *arbitrary*, an essential part of Kierkegaard's view of reality. The author considered that logic does not allow the arbitrary to reach 'inside,' in the midst of the logical system. On the other hand, by constraining reality inside this system, logic doesn't manage to do more than incorporating it, "it has anticipated what it ought merely to predispose" (Kierkegaard 1957, 9). Put differently, if Kierkegaard sees the arbitrary as an essential component of reality, of the real existence, for Hegel, in logic all things happen from necessity and this relation also determines reality. From Kierkegaard's point of view, Hegel conciliated reality and logic, without great success, through the act of mediation. The problem consists in the confusing result:

> Mediation is equivocal, for it designs at once the relation between the two terms and the result of the relation, that in which they stand related to one another as having been brought into relationship; it designates motion, but at the same time rest (Kierkegaard 1957, 11).

Negation is the force that set Hegelian logic in motion. However, Kierkegaard does not regard negation as the catalyst of a proper movement, rather of an immanent motion. For Kierkegaard, any movement has to be transcendental. This implies that logic has to overcome the limits of its own category to reach another. Inside existence, this movement is represented by becoming, which is the transition that the self makes from one stage to another. As a matter of fact, according to Kierkegaard, motion (resembling the meaning of the Greek term κίνησις) is present only in existence, not in the realm of speculation, of the abstract thought (Elrod 1975, 55). In *Philosophical Fragments*, one can read the following:

> Can the necessary come into existence? Coming into existence is a change, but the necessary cannot be changed, since it always relates itself to itself and relates itself to itself in the same way. [...] Everything which comes into

> existence proves precisely by coming into existence that it is not necessary for the only thing which cannot come into existence is the necessary, because the necessary is. […] Nothing whatever exists because it is necessary, but the necessary exists because it is necessary or because the necessary is (Kierkegaard 1962, 91-92).

Motion, with the meaning provided by Kierkegaard, can be understood as becoming, that takes place in the midst of existence, a space destined for coming into being, not just for the simple living. Within this conception, logic is the space of the necessary, to the degree to which the necessary simply *is*, while existence is the space of becoming, to the degree to which the individual self *becomes* (not only is).

However, logic is not the environment of becoming and becoming requires a transcendental motion (at least given the fact that the reason and goal of becoming is represented by the immediate relation with God). In its characteristic synthetic and incisive style, Kierkegaard presents the failed, inadequate relation between logic and motion, in the following fragment from *The Concept of Dread*:

> In logic no movement can come about, for logic is, and everything logical simply is; and this impotence of logic is the transition to the sphere of becoming where existence and reality appear. So when logic is absorbed in the concretion of the categories it is constantly the same that it was from the beginning. In logic every movement (if for an instant one would use this expression) is an immanent movement, which in a deeper sense is no movement, as one easily convinces oneself if one reflects that the very concept of movement is a transcendence which can find no place in logic. The negative then is the immanence of movement, it is the vanishing factor, the thing that is annulled (aufgehoben). If everything comes to pass in that way, then nothing comes to pass, and the negative becomes a phantom. (Kierkegaard 1957, 47).

We can speak of transcendental movement only in the case of a transition towards real existence. However, we know that, according to Kierkegaard, reality cannot be regarded as an (ultimate) part of logic. Nevertheless, Kierkegaard does not advocate the idea that logic lacks motion, but only the fact that this movement remains an immanent one. The negation, the catalyst of movement in logic, represents the immanence of movement or the immanent movement, only that this cannot be a movement in the proper sense (transcendental, according to Kierkegaard), because negation not only negates, not only denies, but also generates opposition. Negation annihilates, but its result is a counter position, not an object that no longer exists. In other words, Kierkegaard points to the fact that negation, in the Hegelian logic, is not a transcendental movement (as it should be, if logic could embed reality), but an immanent one, that does not become something else, that does not cross into existence. Negation remains inside the space of the logical system, a system that only exists. Becoming in logic, if one can use this formula, is achievable inside its

system. The denied object, thus lifted to a superior status, does not transcend the limits of the logical system.

Pointing to the relation between movement and transcendence, Constantin's words gain a much simpler meaning:

> Modern philosophy makes no movement; as a rule it makes only a commotion, and if it makes any movement at all, it is always within immanence, whereas repetition is and remains a transcendence (Kierkegaard 1983, 186).

The comparative analysis of the three concepts presented here, recollection, repetition and mediation, leads to the conclusion that, according to Kierkegaard, only repetition is an authentic movement, one that allows a transcendental motion.[8]

To support this working hypothesis, it could be interesting to take a look upon a lesser known text of Kierkegaard, published in the Supplement of the Hong & Hong edition of *Repetition*. The text is named *Open letter to Professor Heiberg* and is Constantin Constantius' response to a review of the *Repetition*, written by a known contemporary Hegelian Danish thinker, Johan Ludvig Heiberg. In this letter, Constantin raises the problem of repetition in logic:

> There they have called repetition 'mediation.' But movement is a concept that logic simply cannot support. Mediation, therefore, must be understood in relation to immanence. Thus understood, mediation may not again be used at all in the sphere of freedom, where the subsequent always emerges- by virtue not of an immanence but of transcendence. Therefore, the word 'mediation' has contributed to a misunderstanding in logic, because it allowed a concept of movement to be attached to it. In the sphere of freedom, the word 'mediation' has again done damage, because, coming from logic, it helped to make the transcendence of movement illusory. In order to prevent this error or this dubious compromise between the logical and freedom, I have thought that 'repetition' could be used in the sphere of freedom (Kierkegard 1983, 308).

One can notice here Constantin's intention to underline the opposition between immanence and transcendence, with the goal of clarifying the concept of movement, a concept that is crucial for understanding repetition. What is new here is the suggestion that Hegelian philosophy transfers mediation from the space of logic into the space of liberty (existence), a fact that determines the illusory character of the transcendence of motion. For Kierkegaard, repetition is a subjective, spiritual movement that depends on the reality and becoming of the self, aspects that maintain its transcendence. The space of liberty is the space of subjectivity. Liberty is the reality of the self, the domain of self-choice, of decision.

[8] However, Edward Mooney argues that mediation and recollection appear in the text as alternate solutions to the problem of transition or motion, especially of the transition as development, as self-actualization. For instance, it regards the way an individual manages to pass from an aesthetic to an ethical life stage. (Mooney 2007, 286).

I will attempt the transition towards the second kind of inquiry on repetition through the extremely thorough observation offered by Wilson Dickinson. From his point of view, Constantin Constantius (as an author) does not oppose a doctrine of repetition to the notions of recollection and mediation. The possibility of repetition is followed according to the different modes of existence and forms of subjectivity. Recollection, for instance, is not opposed to a theory of repetition that describes reality more accurately, but is opposed rather to an active and contextualized mode of thinking. Then, the definition of the term 'repetition' can be found only in brief descriptions, as the personal reflections of the two characters in the book (Dickinson 2011, 8).

Consequently, the attempt of clarifying Kierkegaard's thought solely by conceptual means is not the best approach of the Danish philosopher's texts.[9] Hence, in the following pages, I will discuss the category of repetition through the perspective of the existential example offered by the two (perhaps three, if we choose to include Job's story) characters analyzed by Kierkegaard for this purpose.

I will thus start with the form of repetition experimented and proposed by Constantin Constantius, the pseudonymous author and central character of *Repetition*.

3. Constantin Constantius or the Failed Act of Repetition

Constantin Constantius made clear his intentions regarding repetition right from the beginning of his writing. He wishes to know whether such a temporal motion is possible and meaningful. Initially, Constantin attached equal importance to repetition and recollection, considering that existence can be meaningful only by accessing one of the two ways of knowing: "If one does not have the category of recollection or of repetition, all life dissolves into an empty, meaningless noise" (Kierkegaard 1983, 149). However, Constantin's aesthetic vision (he is an exponent of this stage) does not allow recollection to become a workable solution for him. A decisive sign of Constantin's aesthetic way of life is his anchoring in the present moment, regarded as discontinuous and perpetually ephemeral.

Constantin considers that recollection requires determining an event or moment with capital existential importance in an already past moment. This generates a temporal separation between the subject and his existential ideal,

[9] In the same article, Wilson Dickinson shows that through the anecdote with Diogenes and the Eeleatics, Kierkegaard sought to bring his reader's attention to the proposed philosophical concepts and philosophical stance that could be identified in his text, but also to the narrative itself, which would point to the roles, expectancies and activity of the reader. The purpose of *Repetition* would become an ethical one, through the empathic involvement of the reader, that brings the text in the practical and interpersonal realm (that pertains to the ethical domain) (Dickinson 2011, 3).

located somewhere in the past. The attitude of the person for whom the remembrance gives life meaning could be summed as follows:

> That is all right, I do not want a relationship with it [n. ideal] now; I just want to remember how it was when I found it (Morris 1993, 311).

So, recollection starts 'by default' with a loss, with an absence, because it does not offer a connection with one's own ideal, established in a past moment and its reactivation in the present. Constantin finds it unacceptable to approach knowledge and existence other than immediately, in a concrete form, right at this very present moment. Thus, if life happens to hold any meaning for him, then it can only be offered by repetition, not recollection.

While reaching these conclusions, Constantin suspected repetition of holding the primary role in re-actualizing that past idealized moment, thus offering its meaning also in the present:

> Indeed, what would life be if there were no repetition? Who could want to be a tablet on which time writes something new every instant or to be a memorial volume of the past? (Kierkegaard 1983, 132-133).

Therefore, Constantin Constantius plans his second travel to Berlin intending to attest the meaning offered by the first trip. Hence he wishes to maintain a present connection with the first idealized image of the trip.

Constantin's vision and solution can be expressed as simply as possible as follows: instead of maintaining our ideal (in a general understanding) in the past, we have to project it into the future, so that we can experience it in the present. The present moment represents the goal from the point of view of experiencing temporality.

Thus, the aim here is to understand how exactly Constantin expects to achieve repetition. Motivated by the perspective of confirming the meaning of existence determined by repetition, he goes to Berlin for the second time, with the intention of reliving the experiences he had a few months ago. As a result, Constantin chooses the same apartment, visits the same café's and attends the same theatre plays, exclusively motivated by the curiosity regarding the possibility of repetition. Unfortunately, the flat that was presented in some earlier paragraphs as one of the most pleasant in Berlin, now appeared gloomy. The most appreciated café from the last visit was not so enjoyable this time and the restaurant with a so hospitable ambiance was now offering monotony, a pathetic twinge of the possibility of repetition.

Neither Königstäter, appreciated in the past for the comic of the farces and for the artistic qualities of its actors, did not generate the same joys and aesthetical satisfactions. The conclusion Constantin reached after all these experiences is at least a strange one:

> The only repetition was the impossibility of a repetition […]. When this had repeated itself several days, I became so furious, so weary of the repetition that I decided to return home. My discovery was not significant, and yet it was

> curious, for I had discovered that there simply is no repetition and had verified it by having it repeated in every possible way (Kierkegaard 1983, 170-171).

This conclusion is reinforced by the surprise he experiences returning home, where, in his absence, his valet turned his house upside down, under the excuse of making a clean sweep of the house. This temporary, although reversible trouble of his intimate space, will conduce Constantin Constantius to the pessimistic conclusion that "I perceived that there is no repetition, and my earlier conception of life was victorious" (Kierkegaard 1983, 171).

I briefly stated before that Constantin Constantius' concept of life is mainly aesthetic. There is no need to review here all the main elements of the conceptions on aesthetic life in Kierkegaard. However, what I'm interested in, at this point, is the aspects, the principles applied to one's own existence, that were 'put in parentheses' by Constantin when he decided to take the trip to Berlin. Constantin's confession from the following paragraph is destined to verify these very aspects:

> ...for I am completely convinced that if I had gone abroad with the idea of assuring myself of it, I would have amused myself immensely with the very same thing. Why is it that I cannot stay within the ordinary, that I insist on principles, that I cannot go around dressed like others, that I like to walk in stiff boots! Do not all agree – both ecclesiastical and secular speakers, both poets and prose writers, both skippers and undertakers, both heroes and cowards- do they not all agree that life is a stream. How can one get such a foolish idea, and, still more foolish, how can one want to make a principle of it (Kierkegaard 1983, 174).

Constantin Constantius' failure is generated by the way in which he decided to test the possibility of repetition that distanced him by the patterns of his ordinary existence. Put differently, Constantin did not receive things as they usually arrive in life, immediately, rather he attempted an intentional experimental approach, theoretically grounded, but with an edificatory existential projection of the effects. By aiming to verify the validity of so-called existential principles, Constantin Constantius reached the conclusion that such a proposal is useless. Principles don't have to dictate your intended purposes; they have to be directly connected to the real existence, immersed in the present, in 'aesthetical purity' but without projecting expectancies (Morris 1993, 334). At this point, Constantin Constantius' view of repetition can be condensed as follows: if repetition is possible, either it cannot be grasped, or it emerges in a rather surprising and unexpected way, within the natural course of ordinary life events.

Constantin Constantius initially discussed repetition with enthusiasm and interest. He now faces disappointment while experiencing the possibility of reaching it. The contrast is total and the conclusion is necessary: *Constantin understands repetition objectively, externally, theoretically, when the success of a*

repetition depends mainly on an authentic personal involvement. He can describe the movement of repetition, but cannot create it.[10]

His attitude is detached as he approaches repetition as a hypothesis that can be empirically confirmed. Moreover, according to a previously cited fragment, although Constantin discovers that repetition is impossible, it does not hold great importance to him. The effects of repetition are neither devastating, nor revealing for his existence. The only thing Constantius gets to confirm through his unsuccessful trip to Berlin is his old attitude on life, a fact that confirms Clare Carlisle's observation:

> Constantin's journey to Berlin is in fact an empty parody of repetition. His failure indicates that he has searched for repetition in the wrong place, and in the wrong way (Carlisle 2005b, 79).[11]

The failure of the Berliner experiment was not caused by the trip itself. As I have already mentioned, not even the intimate space of his own house grants Constantin the possibility of repetition. He is stuck with a paradox, if we take into consideration his position on the movement of recollection. Constantin is looking for a repetition by reproducing an idealized past. The effect generated by this attempt to re-experience the events lived in the past is not the expected one. Events are not refreshed by their re-living, they are rather suffocated by their memory, which can be regarded as a previously established personal norm. At the same time, Constantin becomes aware of his inability to transcend the effects of recollection, when, on his way out, he manages to realize that:

> My home had become dismal to me simply because it was a repetition of the wrong kind. My mind was sterile, my troubled imagination constantly conjured up tantalizingly attractive recollections of how the ideas had presented themselves the last time, and tares of their recollections choked out every thought at birth (Kierkegaard 1983, 169).

Finally, it becomes obvious that Constantin is more capable of recollection, rather than repetition. He had fixed an idealized image in the past to be repeated in an aesthetic frame (as it is, for instance, a theatre hall), in the present. This image becomes the measure of the attempt to test the present external reality (Gouwens 1993, 291). What Constantin Constantius could not reach was the very experience he attempted within repetition and which he amended in the domain of recollection. Maintaining idealized pictures of events from and in a past moment does not go beyond recollection.

The effect of this conclusion practically denies any possibility of repetition, given the fact that Constantin only attempts to reenact a past moment. The object of repetition is not an exterior one. Not the trip is to be repeated, but the remembered joy of that trip, in other words, an 'object' of inwardness, and the remembered joy cannot be subjected to the repetition without experiencing the

[10] From this point of view, Constantin Constantius can be seen mostly as an ironist.
[11] Moreover, the fact that the only plays he attended in Berlin were farces is hardly accidental.

new. However, novelty does not arise from the diversification of the external objects of life, but from the interior transformations required by the movement of repetition.

By proceeding in this way, Constantin neglects his inwardness, his self. By experimentally and externally approaching repetition, he becomes a simple subject within a general test that could attest or infirm the possibility of repetition for anyone, but to a lesser degree for the particular person, for himself. In this way, repetition becomes exterior to the personal edificatory experience.

In the same letter addressed to Heiberg, Constantin Constantius (Kierkegaard) confesses that repetition is just a parody, a fact that generates the confusion about this term:

> The confusion consists in this: the most interior problem of the possibility of repetition is expressed externally, as if repetition, if were possible, were to be found outside the individual when in fact it must be found within the individual. For which reason the young man does indeed do just the opposite, conducts himself quite calmly. The consequence of the journey is that I despair of the possibility and step aside for the young man, who by means of his religious primitivity is going to discover repetition (Kierkegaard 1983, 304).

A summary of the reasons why Constantin fails in his attempt to achieve repetition is necessary. In the first place, the success of a repetition does not consist in validating an all-purpose universal method. Constantin approaches repetition as an experiment, motivated by (an aesthetic) curiosity, rather than animated by the need of internal transformation. In a broader understanding, Kierkegaard does not consider human accomplishment to be determined by the comprehensive understanding of intellectual contrasts, such as that between recollection or repetition and neither by deciding the superiority of one of the two. Accomplishment and human flourishing 'are nourished' by the direct, concrete confrontation with what is problematic for human existence (Mooney 2007, 301).

Secondly, Constantin's failure resulted from his attempt at repetition, being in tune to rather external aspects of existence than to his own subjectivity. Preoccupied by exteriority, Constantin is an observer of his own existence and of his intellectual projections (Carlisle 2005a, 531). This generates his difficulty of getting involved in a relationship with the surrounding reality of any kind, other than conceptually and ideally. Also, Constantin reveals this aspect of his personality on at least two separate occasions:

> ...as a rule I tend to relate to men as an observer (Kierkegaard 1983, 134).

> This, then, is the thanks one gets for having trained oneself every day for years to have only an objective theoretical interest in people and also, if possible, in everyone for whom the idea is in motion! (Kierkegaard 1982, 180).

The third and maybe the most important reason for Constantin's failure could be questioningly expressed in the following way: is repetition something a person could reach only through his volitive efforts of any kind? Does the success of repetition depend on the individual? For the moment, I will limit my answers here to negative responses to these questions, emphasizing the fact that Constantin's repetition is a project initiated by himself, whose success is conditioned by his own will and unfolded along some pre-established steps. Exclusively conceived as a human endeavor or as the effect of reasonable expectations, repetition is impossible. As we will see in the case of Job, the essence of repetition consists in assuming the shock of its impossibility, as human finality (Mooney 1998, 289-290, 300).

Constantin's understanding of repetition stops with the failure of his experiment. His conclusion is that in life you can never be truly satisfied, an attitude that Constantin assumes, together with the aesthetic self-sufficiency that recognizes the impossibility of repetition (Gouwens 1993, 291). He accepts his failure, this gesture being the most important existential step he managed to make because he realized that repetition is a transcendental, religious movement (Carlisle 2005b, 79). Viewed from this perspective, Constantin's failure gains the nuances that open the possibility of investigating other forms of repetition, that of the young lover and Job's. I will first take care of the young lover's case.

4. Repetition of the self

The presence of the young man within the pages of the *Repetition* is marked by three progressively linked moments that determine the process of becoming and the inner transformation. These three moments could be expressed schematically as follows:

1. Moment one: the confessional relation, friendship with Constantin Constantius;
2. Moment two: putting his own existence under the example of Job's trials;
3. Moment three: the re-turn, the re-discovery of the self as an expression of repetition;

1. The first part of the *Repetition* conveys the apparition, in a rather secondary role, of the young man, who will not express himself directly until the second part of the book, through the letters addressed to Constantin. Initially, we get acquainted with the character through the stories and observations made by Constantin Constantius, regarding a love story confessed by a young, melancholic man, that is under the influence of the friendship and advice he gives. As Clare Carlisle also mentioned, the first thing we know about this young man is his predisposition towards melancholy, a feature that is specific to aestheticism, under the form of indifference and immobility, as existential attitudes (Carlisle 2005a, 533). The young man's melancholy determines his preoccupation not only with the joy of loving the girl with whom he has fallen in love, but rather

with the avenue of losing her. This observation shifts our attention towards the comparison between recollection and repetition. Through Constantin's voice, we find out that the young man experiences his love by putting it through the lenses of recollection, a fact that determines, on one hand, the lack of happiness in his relationship with the girl and, on the other hand, his impossibility of believing and understanding repetition:

> He was deeply and fervently in love, that was clear, and yet a few days later he was able to recollect his love. […] Recollection has the great advantage that it begins with the loss (Kierkegaard 1983, 136).

> My young friend did not understand repetition; he did not believe in it and did not powerfully will it. […] If the young man had believed in repetition, what great things might have come from him, what inwardness he might have achieved in this life! (Kierkegaard 1983, 145-146).

Constantin finds it problematic that his young friend cannot experience his love fully and concretely, because of his tendency to idealize it. As noticed before, this idealization is a movement of recollection. Suspending a moment in a past time, with the intention of establishing an existential reference, inhibits its actualization in the present. From Constantin's perspective, the condition of the young lover perfectly fits this model. The young man idealizes his love, otherwise sincere and intense, a love that, in fact, he cannot enjoy anymore. The young man's dilemma and suffering are based on the unhappy remembrance of what he should actually hope to encounter.

At this point, it is necessary to highlight the difference between Constantin's and the young man's perspectives. While for Constantin, the negative conclusion about recollection brings him to an experiment based on external criteria for repetition, for the young man, the experience of love as recollection becomes an occasion for inner (re)discovery.[12]

It is worth mentioning here Stephen Crites' observation: the term *recalling*, more exactly, its Danish correspondent, *Erindring* (in German, *Erinnerung*), literally means interiorization (Crites 1993, 232). The fact that the young man falls in love does not materialize the relation with the girl, as one might expect. The incident puts him into a rather poetic disposition, maintained by the constant longing for his lover. Love is transformed into longing, wistfulness. The young man discovers an absence in himself, a void that, to his despair, cannot be replaced by the real presence of his lover. This is what Constantin Constantius wrote about this:

> Nevertheless, he did not still love her, because he only longed for her [...]. The young girl was not his beloved: she was the occasion that awakened the poetic in him and made him a poet. That was why he could love only her, never forget

[12] In the same supplement, we can read the following: "His being has been split, and so it is not a question of the repetition of something external but of the repetition of his freedom" (Kierkegaard 1983, 304).

> her, never want to love her, and yet continually only long for her [...]. As time went on, his state became more and more anguished. His depression became more and more dominant (Kierkegaard 1983, 137-138).

When the effects of an event don't fit the expectations, we face a dilemma and this dilemma requires a decision. The young man's dilemma is, in this case, an ethical one. On the one hand, he feels responsible for the girl he fell in love with and to whom he confessed his love and, on the other hand, he feels responsible towards his own self, who pushes him into a solitary existence, determined by his melancholic nature. For this young man, the girl has a tremendous importance, but not as herself, but through the effect she has on him "...The girl was not an actuality, but a reflection of motions within him and an incitement of them" (Kierkegaard 1983, 185).

Preoccupied with all these thoughts, the young man turns his hopes towards his confessor, Constantin, and his advice.

One can ask oneself whether Constantin Constantius' conclusions and observations regarding the young man could be taken seriously, as his own experiment on the possibility of repetition failed. It is true that he failed to experience repetition, but this does not mean that, conceptually, Constantin didn't apprehend this existential movement. What he reproached to his young friend is the fact that *he didn't understand repetition, not that he couldn't experience it*. Constantin accepted the fact that, for him, repetition is a transcendental, religious, thus inaccessible movement, a movement that is considered, however, by the young man, the solution to his own existential dilemmas. By moving from the understanding of repetition in the absence of the ability to experience it, to the need of living it even without a full understanding, the young man moves from the influence Constantin had on him to the one manifested by Job's example.

The first signs of this mutation were already visible when the young man refused to accept Constantin's plan to liberate him from the burden of his love. As I already mentioned earlier, the young man's dilemma is an ethical one, a dimension that Constantin cannot grasp, his advice being based on his personal, failed repetition experiment. Consequently, Constantin's plan according to which the young man could break the engagement with the girl, exiting the relation clean-handed, proves to be an unacceptable solution for the honest young man; this is also due to his lack of 'ironic resiliency' as Constantin labeled it (Kierkegaard 1983, 137). The young man materializes his refuse and, together with it, the end of the confidential, direct relation with Constantin, embarking on a solitary trip to Stockholm. This transfer is captured by Constantin with the lucidity of a reflexive person, detached from the 'thunders' of life:

> The issue that brings him to a halt is nothing more nor less than repetition [...]. It is fortunate that he does not seek any explanation from me, for I have abandoned my theory, I am adrift. Then, too, repetition is too transcendent for me. I can circumnavigate myself, but I cannot rise above myself. I cannot find

> the Archimedean point. Fortunately, my friend is not looking for clarification from any world-famous philosopher or any professor publicus ordinarius [regularly appointed state professor]; he turns to an unprofessional thinker who once possessed the world's glories but later withdrew from life – in other words, he falls back to Job (Kierkegaard 1983, 186).

The young man's distancing from Constantin symbolized the transformation that took place in his own conscience, translated as the transition from the ideality to reality, from philosophy to existence, an idea that, also delineates the concept of repetition and which is expressed, in a metaphorical sense, in the first passage of *Repetition* (the example of the dispute between Diogenes and the Eleatics) (Carlisle 2005a, 534). I believe that the relation between Constantin Constantius and the young man personified the opposition between recollection and repetition. More specifically, Constantin intended to experience repetition and, in turn, he achieved recollection; the young man is suffering through recollection and projects his salvation through the movement of repetition.

2. The young man approaches Job with the hope of achieving repetition. For this, he decides to embark on a trip to Stockholm, distancing in this way from his lover and from Constantin's dubious influence. With these events, Kierkegaard changes the narrative register of the *Repetition* that marks the inner transformations of the young man. The writing pen passes from Constantin to the young lover, who signs eight confessional letters that were surprisingly addressed to the same Constantin. Although he recognizes his malign influence, especially when he was constantly in his company, the young man could not repress a certain attraction towards his ambiguous personality:

> You hold me captive with a strange power. There is something indescribably salutary and alleviating in talking with you, for it seems as if one were talking with oneself or with an idea. Then, upon finishing speaking and finding solace in this speaking out, when one suddenly looks at your impassive face and reflects that this is a human being standing before one (Kierkegaard 1983, 188).

Actually, the one who manifested ambiguity was the young man himself. Firstly, he wasn't provided a name, as Constantin Constantius was. His personality swung between Constantin's influence and Job's existential example; he was thus a character that wasn't fully defined yet. Along the entire length of *Repetition*, the young man experiences a process of becoming, at the end of which we expected him to adopt one of the two alternatives of being or, even a different one, generated by his own decisions, meditations and experiences. In any case, the young man is a character defined by many open possibilities.

The young man explicitly expressed his admiration for Job in the second letter sent to Constantin. His admiration comes with his wish to tame his own suffering; not only by accessing a theoretical, philosophical solution, "what

miserable worldly wisdom poorly affords" (Kierkegaard 1983, 198) but by the power of a concrete example of an existential circumstance.

I have shown before that the reason the young man has chosen Job as his 'hero' was Job's successful repetition. From the young man's perspective, Job accomplishes repetition when, after losing everything, he then regains everything twofold. The young man validates the intensity of his suffering from losing his lover by comparing it to those experienced by Job:

> I have not owned the world, have not had seven sons and three daughters. But one who owned very little may indeed also have lost everything; one who lost the beloved has in a sense lost sons and daughters, and one who lost honor and pride and along with it the vitality and meaning of life – he, too, has in a sense been stricken with malignant sores (Kierkegaard 1983, 198-199).

Therefore, suffering becomes the category through which the young man hopes to get closer to Job. Suffering distances the young man from Constantin's influence and brings him closer to Job's lesson. I reassert the idea that the repetition proposed by Constantin is only an experiment, a reflexive game, that doesn't require authenticity and personal involvement. Suffering, in contrast, involves a higher level of interiorization of the lived experiences and the young man is undoubtedly suffering, similarly to Job. The young man finds himself in the middle of a symmetric paradox, whose central element is suffering. On the one hand, Constantin's detached reflection limits his understanding of the very thing upon which he reflects the most, suffering. On the other hand, Job and all those who followed his example – 'proposed' a type of patient, wordless suffering, allowing the full understanding of the process, by avoiding the detached reflection but sharing the situation of the sufferer (Burges 1993, 254).[13]

There are two ways of understanding, personified by the two characters of the *Repetition*, Constantin Constantius and his young friend. As I have already asserted, Constantin's understanding of an idea, a concept or a person is objective, supported by reflection and his critical skills. His understanding is observational, discursive, detached. On the other hand, the young man (separated from Constantin's influence) develops a subjective kind of understanding, to the degree to which his actions are determined by personal involvement, self-conscience and affective intensity. Constantin wishes to find out whether repetition is possible, his motivation being only a simple experimental curiosity. The young man wishes to experience repetition to cut his suffering and inner turbulence generated by his love story. His intention is not to understand repetition by experiencing it, *with the goal of understanding it*, rather he wants to live and understand it simultaneously, without one action causing the other. In the case of the young man, one can speak of coherence between the

[13] I owe Andrew Burges this view of paradox, that places Job and Constantin in opposition, a vision that was developed in a study that proposes two central themes for the *Repetition*: repetition, as Constantin's perspective and suffering, as the young man's perspective.

concepts and his own existence. A category such as repetition does not represent a simple intellectual curiosity, but an existential aim.

With such an openness of the soul, the young man approached Job. Besides, in the first passage of one of the *Edifying Discourses* signed by Kierkegaard and entitled *The Lord Gave, and the Lord Hath Taken Away, Blessed Be the Name of the Lord,* we learn about the moral of Job's parable:

> Not only do we call that man, a teacher who through some particularly happy talent discovered, or by unremitting toil and continued perseverance brought to light one or another truth; left what he had acquired as a principle of knowledge, which the following generations strove to understand, and through this understanding to appropriate to themselves. Perhaps, in an even stricter sense, we also call that one a teacher of men who had no doctrine to pass on to others, but who merely left himself as a pattern to succeeding generations, his life as a principle of guidance to every man, his name as an assurance to the many, his own deeds as an encouragement to the striving. Such a teacher and guide of men was Job, whose significance is by no means due to what he said but to what he did. (Kierkegaard 1958, 67).

Job's lesson is not a dogmatic one and the lack of this feature determines the young man's empathy in assuming his words:

> I do not read him as one reads another book, with the eyes, but I lay the book, as it were, on my heart and read it with the eyes of the heart, in a clairvoyance interpreting the specific points in the most diverse ways. […] Every word by him is food and clothing and healing for my wretched soul (Kierkegaard 1983, 204).

At this point, it is necessary to return to a term that I previously emphasized in the pages dedicated to Constantin's condition. Returning from his trip to Berlin, he recognized that "repetition is too transcendent for me. I can circumnavigate myself, but I cannot rise above myself" (Kierkegaard 1983, 186). Transcendence, from Constantin's perspective, is the capacity to overcome certain limitations, here the limit being set by one's self. His existence is immanent, given the fact that Constantin does not cross the pre-established limits. Constantin's limit is represented by its reflexive and objective way of knowing, a limit that the young man surpasses, at least intentionally, by comparing himself, from the point of view of repetition, to Job. What the young man identified as a transcendent category in Job are the trials he is subjected to, the *ordeal*. "This category, ordeal, is not esthetic, ethical, or dogmatic – it is altogether transcendent." (Kierkegaard 1983, 210).

The ordeal transcends knowledge and requires suffering to be understood. Thus, it becomes obvious that the ordeal cannot be understood as an aesthetic category. Constantin's experimental repetition and everything I have concluded about him until this moment seem to advocate this idea.

The ordeal can neither be an ethical category, because the trials to which Job was subjected were not the result of divine punishment or (the sense of?)

guilt. All the rational, human explanations (see the interventions of Job's friends) lead to the idea, even if unconscious, of Job's guilt:

> To him every human interpretation is only a misconception [...] It can be very becoming and true and humble if a person believes that misfortune has struck him because of his sins, but this belief may also be the case because he vaguely conceives of God as a tyrant, something he meaninglessly expresses by promptly placing him under ethical determinants (Kierkekaard 1983, 207).

The ordeal could be dogmatic only to the degree to which there could be a science that integrates it within its system of justifications. Such a validation of the ordeal could not be possible because "as soon as the knowledge enters, the resilience of the ordeal is impaired, and the category is actually another category" (Kierkegaard 1983, 210).

A category like the ordeal cannot have but a personal stake and validity, subjected to the human as an individual. The young man understands that the experience of the ordeal situates Job into a direct relationship with God and such a relation excludes the validity or the possibility of 'any explanation at second hand' (Kierkegaard 1983, 210) (thus, an ethical, dogmatic, or theoretical explanation of any kind). Entirely convinced by the possibility of a repetition, the young man identifies the very moment when this could have occurred in Job:

> When every *thinkable* human certainty and probability were impossible. Bit by bit he loses everything, and hope thereby gradually vanishes, inasmuch as actuality, far from being placated, rather lodges stronger allegations against him (Kierkegaard 1983, 212).

Reaching this point, the two characters parted. The young man wishes to regain his lover, not by searching for God; also, his suffering does not overcome any limit of human understanding and empathy. Thus, our hero finds himself in the situation of having to turn his attention inwards.

3. The young man initially believed that the motivation to become a good husband could be identified with a successful repetition:

> I am waiting for the thunderstorm – and for repetition. [...] What will be the effect of this thunderstorm? It will make me fit to be a husband. It will shatter my whole personality – I am prepared. It will render me almost unrecognizable to myself (Kierkegaard 1983, 214).

From this perspective, the young man faces a new series of issues and dilemmas, now on an ethical dimension, in the same time, provoking a disturbance of his own inwardness. The expression of the intensity of this turbulence is captured by the thunder metaphor. The thunder represents the event that could make repetition possible, an event that could start his interior transformation and could orient his subsequent existence. In Job's case, the

young man identifies the thunder with God's 'initiative' to test him, something that causes the possibility of repetition.[14]

Going back to the young man, the above mentioned aspects stress, once again, the dualism and ambiguity of his personality. What the young man is striving to repeat is still unclear. We cannot understand what he wishes, whether it is the re-actualization of the relation with the beloved girl or her forgetting, in order to become once again free and capable of self-discovery. A situation that opens the possibility of a decision and thus the movement of the young man towards the ethical stage of existence (either he chooses marriage or himself) is solved by an external event. The young man's 'thunder' is represented by the news of the marriage of his former lover. This event naturally determines a return to his inner self and a process of self-discovery. The news brings him the opportunity of experiencing repetition in a different form from his first expectations:

> She is married – to whom I do not know, for when I read it in the newspaper I was so stunned that I dropped the paper and have not had the patience since then to check in detail. I am myself again. Here I have repetition; I understand everything, and life seems more beautiful to me than ever. It did indeed come like a thunderstorm, although I am indebted to her generosity for its coming (Kierkegaard 1983, 220).

Repetition is achieved accidentally; at least this is the interpretation the young man attributed to regaining his self:

> Is there not, then, a repetition? Did I not get everything double? Did I not get myself again and precisely in such a way that I might have a double sense of its meaning? (Kierkegaard 1983, 220-221).

Freed from the moral pressure and self-generated inertia, the young man (re) sets his self in motion. He understands and accepts the fact that the girl was nothing but an occasion for self-rediscovery: "it also gave me what I loved more – myself, and gave it to me through generosity" (Kierkegaard 1983, 220). His engagement determined an unpredictable change of self and his impossibility to decide brought him to a state of immobility. An external event resets his self in motion and his existence regains meaning. The young man is freed from the past by redefining it and this is how he regains the importance of the present moment.

In the aforementioned supplement, Constantin Constantius emphasizes:

> In the explanatory letter it says, 'The young man explains it as the raising of his consciousness to the second power.' This certainly ought to be the most definite

[14] One can observe here a similarity with the position of another biblical character, Abraham. The possibility of belief/ repetition by the virtue of the absurd is conditioned by the attempt or the trial designed by God, to which the person reacted. Faith and repetition are not movements of the self that could be exclusively generated by human possibility.

> expression of the fact that I conceive of repetition as a development, for consciousness raised to its second power is indeed no meaningless repetition, but a repetition of such a nature that the new has absolute significance in relation to what has gone before, is qualitatively different from it (Kierkegaard 1983, 307).

Unlike Constantin, the young man manages to achieve repetition. He approaches existence not through external realities but through his inwardness. Repetition, unlike recollection, which only amplifies the experience for its clarification or intensification, is thus a movement of self-transformation (Crites 1993, 241).

Not only Constantin Constantius, but also the young man, personify two existential perspectives, in a process of movement (of becoming, in a more pronounced sense, in the case of the young man) along the *Repetition*. The first is a philosopher, whose movement is blocked inside a skeptical attitude that makes the conclusion that repetition is not possible. The second is represented by the lover, who is animated by passion and rediscovers his self and becomes his self because of the ethical crisis generated by the unhappy love story (Carlisle 2005a, 523).

> The cause of the repetition's failure, in the case of Constantin, and the success of the young man is the same: their specific view of knowledge, love, recollection, repetition and existence. Their limits in understanding and living are traced by their own selves. In this case, however, the reached limit is not an obstacle, rather a source of existential stability. Not managing to achieve repetition, Constantin certifies in fact the validity of his earlier view on life. Moreover, one can notice an intensification of his contentment towards the monotony and uniformity of his own existence. In a similar way, the young man's love story produces the satisfaction of regaining his self, as the result of a kind of repetition adapted to his capacities. The limits achieved by the two in the realm of repetition become evidence for the validation of their life perspectives. The experiences they lived lead to a greater awareness of existence.
>
> What I would like to stress here, as a possible version of the *Repetition's* conclusion, is Kierkegaard's intention to emphasize the importance of existential consistency. The way we understand existence, its principles, have to correspond to how we experience it. In other words, the way we live our life has to avoid contradiction with the way we understand it. What happens, though, when, despite such coherence, not only existence, but the capacity to understand, are subjected to a trial? How could we answer with a sense of understanding when our own existence becomes an ordeal? Finally, what is the human attitude that converts paradox into liberty? Through such interrogations, I will turn my research of the category of repetition as it is presented in the text where Kierkegaard discusses *The Book of Job*.

5. The Example of a Genuine Repetition - Job

Although between the pages of *Repetition* we could read about Job's ordeal as an example of authentic repetition, Kierkegaard doesn't develop a comparative

analysis of the biblical character; as he does with Constantin and the young man. A possible cause for this situation could be Kierkegaard's distinction between religious topics (approached in the texts signed and assumed by his real name) and his aesthetical writings (here we include all the books he published under pseudonym). Most probably, the seriousness of a biblical motif and the trials Job is subjected to, determines Kierkegaard to change the register of his writing and to transfer the analysis of repetition in one of the four *Edifying Discourses*, edited in the same year he published *Repetition* and *Fear and Trembling*.

Throughout this work, one is presented with the opportunity to discover, either by the analysis of the failed forms or repetition, or with the help of the textual clues left by Kierkegaard, that repetition, in its genuine form, is a religious category, transcendent, by virtue of the absurd, whose temporal correspondent is everlasting or the eternal. One of the clearest theoretical expressions of the concept of repetition can be found, once again, in *The Concept of Dread*. In a large footnote, the pseudonymous author Virgilius Haufniensis, distinguishes between repetition in the natural world, considered a necessary movement and repetition in the spiritual realm, which requires inwardness. Regarding the latter kind of repetition, Haufniensis reminds us that it emerges on the strength of religion and that it represents a transcendental, religious category, a movement by virtue of the absurd and that eternity is the real repetition.[15]

Kierkegaard does not intend to show directly, within the discourse about Job, this form of repetition. However, the references to Job found in the text of the *Repetition*, and the characteristics of this new philosophical category underlined above, encourages us to approach the Kierkegaardian version of this biblical chapter as the ultimate expression of the concept of repetition.

The first remark made by Kierkegaard is that, in the moment he was stricken, Job reminds himself and others, that first, the Lord was merciful, endowing him with everything that he has now lost. Thus, loss generates his gratitude in the first place. This view gave rise to the thought that, for something to be taken away, it has to be rendered first with generosity; this is why, prior to mourning his loss, Job expressed his gratitude:

> At the moment when the Lord took everything, he did not say first 'The Lord took,' but he said first, 'The Lord gave.' The word is short, but in its brevity it perfectly expresses what it wishes to indicate, that Job's soul is not crushed down in silent submission to sorrow, but that his heart first expanded in gratitude; that the loss of everything first made him thankful to the Lord that

[15] "...the passion of the absurd to which the concept of 'repetition' corresponds. […] for in faith repetition begins. […] In the sphere of nature repetition exists in its immovable necessity. In the sphere of spirit the problem is not to get change out of repetition [...] but the problem is to transform repetition into something inward, into the proper task of freedom [...] Here the finite spirit falls into despair. This Constantine has indicated by stepping aside and letting repetition break forth in the young man by virtue of the religious" (Kierkegaard 1957, 16-17).

> He had given him all the blessings that He now took from him. (Kierkegaard 1958, 75).

Here, we can find a first understanding of the movement of repetition. In Job's case, the possibility of repetition first manifests as a consequence of the initial existence of an object 'to be repeated,' of a 'good' that can be lost. To be repeated, one needs to own a 'good,' a desired object, a beloved person, that one enjoys; and Job shows this attitude by first manifesting his gratitude. Gratitude is something that distinguishes Job from the two central characters of the *Repetition*. Constantin Constantius approaches repetition as a simple, objective and external experiment, while the young man accidentally achieves repetition, even if this is reached at a superior level of inwardness. However, neither of the two characters reacts to the movement of repetition – or through repetition – as a fundamental response, a limit of their existence. Repetition is not an assumed action, but rather objective or haphazard. In Job's case, however, gratefulness demands a deliberate and responsible reaction to a divinely imposed burden; an act, as in Abraham's case, that lacks ethical, rational justification.

Gratefulness indicates the possibility of authentic repetition. Through gratefulness, remembering what he had until then, as effect of God's grace, did not occupy Job's conscience in the form of a tormenting memory. This is in contrast to the consequences of repetition in Constantin's and the young man's case. We are reminded at this point of the disappointing effects of Constantin's visit at Berlin, as an attempt to repetition through recollection, and the inner torment of the young man, troubled by the memory of his love and then finding (him) self, but only in a melancholic and skeptical disposition, as baseline existential attitude.

Beginning with gratefulness, conceived as a right reaction to unfortunate life events, Kierkegaard presents, in contrast, the human, ordinary, rational and justifiable attitudes that usually occur as reactions to ordeal. For instance, we can find among the corresponding chapters, not only the situation, but also Constantin Constantius' and the young man's (let's call them existential) reactions. I illustrate here, with some revealing passages, first the representation of the young man, then Constantin's perspective, as failed attempts to integrate the loss and the memory of the lost 'good;' in other words, the effects of the failed attempts at repetition:

> Thus his happiness became pernicious to him; it was never lost, but only lacking, and it tempted him more in the lack than ever before. What had been the delight of his eyes, he desired to see again, and his ingratitude punished him by conjuring it up as more beautiful than it had formerly been. […] Thus he condemned his soul to living famished in the never satisfied craving of want. (Kierkegaard 1958, 77).

> In fact, who would ever finish, if he wished to speak about what so frequently has happened, and will so frequently be repeated in the world? Would he not tire far sooner than would passion of that ever new ingenuity for transforming

> the explained and the understood into a new disappointment, wherein it deceives itself! (Kierkegaard 1958, 78).

The second expression, "The Lord took away," illustrates Job's gaining of an insight on the reason of his loss. Similar to Abraham, Job recognizes God as the source of his ordeal. Job puts everything in God's hands and, in this way he distances himself from any possibility of providing a human, rational explanation or justification. Job does not blame the thunders, the enemy tribes or human nature for being generally corrupt, but accepts everything as coming from God (with the same attitude with which he manifested gratitude). This orientation towards God is not interpretable as guilt-inducing (a sign of rebellion, indignation or affront). Job decides to bear his ordeal and in doing so he also acknowledges his intimate relation to God. Just like Abraham, Job trusts God by the virtue of the absurd. Job trusts that everything will be returned to him due to a simple but terrible reason: to God, everything is possible. In this way, Job enters the category of the 'exceptional Individual,' because Job

> did not retard his soul and extinguish his spirit in reflections or explanations which only engender and nourish doubt, even if the one who dwells on them does not realize it (Kierkegaard 1958, 81).

Actually, each of the three phrases that underline Job's exceptional character is doubled by the emphasis on the human, ordinary versions of reaction to the ordeal. To quote Johannes de Silentio in *Fear and Trembling*, this attempt underlines a teleological suspension of ethics. One by one, Kierkegaard puts under the questioning mark (and subsidiary, under the sign of failure) all the possible versions of reacting to the possibility of a spiritual trial. As a result of these interrogations, as in the case of Abraham's faith, the paradoxical nature of repetition is revealed. Because, eventually, similarly to Abraham, Job is a father of faith; the act of belief is the existential movement of repeating 'God's suffering in time,' the expression of the Absolute Paradox (Taylor 2000, 258).

It is worth mentioning here that the instances where the term *repetition* appeared in the text of this discourse have a deceiving effect; their aim is, however, to bolster the discrepancy between Job's success and attitude and the failure of the other presented trials. In this discourse, Kierkegaard assembles under this term the variety of the possible human behaviors, as potential reactions to Job's situation. Thus, repetition becomes an empty, hollow and specifically human endeavor, while the true kind of repetition is specific to the exceptional lives.

The last part of Job's discourse highlights God's grace. The ordeal is the event that, ultimately, makes faith possible. God took everything from Job, but not his peace of the heart, the spiritual balance that gave birth to that kind of belief by virtue of the absurd.[16] Peace of heart represents the measure of Job's

[16] "Hence the Lord did not take everything, for He did not take away Job's praise, and his peace of heart, and the sincerity of faith from which it issued; but his confidence in the Lord

faith in God, springing from the certainty of a life without sin. In this way, Job expresses the liberty that is not suppressed by the characteristically human thought of believing that the ordeal and suffering are the effect of previous sins. Job affirms that his thinking is right and, in this way, he reveals his freedom, which is not a simple act of rebellion, as, through it, he chooses to sincerely believe in God and in his own innocence. The ordeal becomes an existential extreme situation, because it offers the person the right perspective on self-actualization: the becoming of the self does not rely exclusively and with certainty in the hands, power and will of the man, but is conditioned by divine intervention (Diaconu 1996, 154).

Job's words show the entire process of the temporal movement of repetition. Firstly, the person is offered everything, after which, he is deprived of it, to finally bring his contribution, through gratefulness and contentment for what he was given, in the very moment he has lost everything. The individual doesn't need to seek for explanations or rationalizations for his suffering, the spiritual trials, the experienced ordeals or joys, in the human goodness or badness, in the good or vicious nature of the human being. Similarly, the success of repetition is not an act that is (exclusively) subjected to will. The person does not *reach repetition*, he only *brings his contribution* to the genuine repetition; his contribution resides in suffering, pains, infinite resignation and belief in the virtue of the absurd, as reactions to the spiritual trial he is subjected to.

References:

Burges, Andrew. 1993. "Repetition – A Story of Suffering." In *International Kierkegaard Commentary. Fear and Trembling and Repetition*, ed. Robert L. Perkins, 247-262. Macon. Georgia: Mercer University Press.

Carlisle, Claire. 2005a. "Kierkegaard's Repetition: The Possibility of Motion." *British Journal for the History of Philosophy* 13 (3): 521-541.

____. 2005b. *Kierkegaard's Philosophy of Becoming*. New York: State University of New York Press.

Crites, Stephen, 1993. "The 'Blissful Security of the Moment:' Recollection, Repetition, and Eternal Recurrence." In *International Kierkegaard Commentary. Fear and Trembling and Repetition*, ed. Robert L. Perkins, 225-246. Macon, Georgia: Mercer University Press.

Diaconu, Mădălina. 1996. *Pe marginea abisului; Søren Kierkegaard și nihilismul secolului XIX.* București: Editura Ştiinţifică.

remained with him as before, perhaps more fervently than before; for now there was nothing at all which could in any way divert his thought from Him" (Kierkegaard 1958, 83).

Dickinson, Wilson. 2011. "Repetition, Not Simply Recollecting, Repetition, On Kierkegaard's Ethical Exercises." *Springer Science+ Business Media B.V.*: 8. Published online: 02 August.

Elrod, John. 1975. *Being and Existence in Kierkegaard's Pseudonymus Works.* Princeton: Princeton University Press.

Gouwens, David J. 1993. "Understanding, Imagination, and Irony in Kierkegaard's Repetition." In *International Kierkegaard Commentary. Fear and Trembling and Repetition*, ed. Robert L. Perkins, 283-308. Macon, Georgia: Mercer University Press.

Kierkegaard, Søren. 1957. *The Concept of Dread*. Princeton: Princeton University Press.

____. 1958. *Edifying Discourses A Selection*. New York: Harper Torchbooks/The Cloister Library.

____. 1962. *Philosophical Fragments*. Princeton, New Jersey: Princeton University Press.

____. 1983. *Fear and Trembling, Repetition*. Princeton, New Jersey: Princeton University Press.

Lowrie, Walter. 1938. *Kierkegaard*. London and New York: Oxford University Press.

Jowett, B. trans. 1892. *The Dialogues of Plato*. London: Humprey Milord: Oxford University Press.

Mooney, Edward F. 1998. "Repetition: Getting the world back." In *The Cambridge Companion to Kierkegaard*, ed. Alastair Hannay and Gordon D. Marino, 282-307. Cambridge: Cambridge University Press.

____. 2007. *On Soren Kierkegaard, Dialog, Polemics, Lost Intimacy and Time.* Hampshire, Ashgate.

Morris, T. F. 1993. "Constantin Constantius's Search for an Acceptable Way of Life." In *International Kierkegaard Commentary. Fear and Trembling and Repetition*, ed. Robert L. Perkins, 309-334. Macon, Georgia: Mercer University Press.

Taylor, Mark C. 2000. *Journeys to Selfhood.* New York: Fordham University Press.

Algebra of Theoretical Term Reductions in the Sciences

Dale Jacquette

Abstract: An elementary algebra identifies conceptual and corresponding applicational limitations in John Kemeny and Paul Oppenheim's (K-O) 1956 model of *theoretical reduction* in the sciences. The K-O model was once widely accepted, at least in spirit, but seems afterward to have been discredited, or in any event superceeded. Today, the K-O reduction model is seldom mentioned, except to clarify when a reduction in the Kemeny-Oppenheim sense is not intended. The present essay takes a fresh look at the basic mathematics of K-O comparative vocabulary theoretical term reductions, from historical and philosophical standpoints, as a contribution to the history of the philosophy of science. The K-O theoretical reduction model qualifies a theory replacement as a successful reduction when preconditions of explanatory adequacy and comparable systematicization are met, and there occur fewer numbers of theoretical terms identified as replicable syntax types in the most economical statement of a theory's putative propositional truths, as compared with the theoretical term count for the theory it replaces. The challenge to the historical model developed here, to help explain its scope and limitations, involves the potential for equivocal theoretical meanings of multiple theoretical term tokens of the same syntactical type.

Keywords: John Kemeny, Kemeny-Oppenheim (K-O) model of theoretical reduction, Paul Oppenheim, theoretical reduction, science, scientific theory

1. Kemeny-Oppenheim (K-O) Model

The reduction of secondary to primary sciences encounters difficulties where reduction procedures are described as involving comparisons of unspecified 'terms' that leave their individuation and denumeration undetermined. The 1956 Kemeny-Oppenheim (K-O) model of scientific and more generally theoretical reduction prescribes a reduction procedure that involves a method for the array and elimination of theoretical 'terms,' but does not explain what is to count as a *term*.

The omission turns out to have important implications in applying the K-O model. The difficulties entailed by this lack of clarity about the nature of terms apply to Kemeny and Oppenheim's treatment of theoretical reduction, but can also be raised in a general way against any attempt to set forth procedures of reduction that involve enumerations of theoretical terms and vocabularies before and after the replacement of one set of equivalently explanatorily capable

© Symposion, 1, 1 (2014): 51-67

competing theories for another at the propositional level. Kemeny and Oppenheim do not suggest that a reduction is achieved when the number of theoretical terms is reduced from alternative equally explanatorily powerful and systematic theory.

When K-O preconditions are satisfied, when the reducing theory explains all the same relevant observational data that the reduced theory explains, and the reducing theory is at least as well systematized as the reduced theory, however the concept of being systematized is more exactly interpreted and applied, then there results a K-O theoretical reduction marked criteriologically by a reduction in the number of theoretical term types from the reduced to the reducing theoretical vocabulary. The further moral in the fate of the K-O model of theoretical reduction in the sciences has to do with the limits of considering only syntax tokens and types, and the need also to go beyond Kemeny and Oppenheim by including the meanings and full-blooded semantic interpretations of terms and expressions in an adequate metatheory of the term token economy in comparative theoretical explanation. There is, in other words, more to reduction, even when K-O preliminary conditions are satisfied, than counting up the number of minimally needed term tokens on both sides of a theoretical reduction undertaken at the propositional level, when a reduced theory is replaced by a reducing theory.

The informal discussion Kemeny and Oppenheim present in their influential co-authored essay "On Reduction" explains theoretical reduction in the sciences in terms of several factors. When satisfied, they are supposed to produce as a consequence a numerical reduction in the number of theoretical terms needed to express the truths of reduced and reducing scientific theory for purposes of comparing their respective cardinalities. The comparative vocabulary K-O theoretical reduction model, as the authors acknowledge and intend, is easily and equally attractively extended to all systematic branches of knowledge possessing an identifiable terminology in which explanations are expressed. It is accordingly not just our understanding and ability intelligently to pursue theoretical reductions within the natural sciences in a narrow sense that is at stake, but all propositional knowledge involving explanatory propositions. Kemeny and Oppenheim believe that theoretical reduction contributes to progress in scientific understanding, because it brings science closer to more basic principles of explanation, which can in turn make a scientific theoretical explanation more practically applicable, easier to grasp in its most fundamental principles, and potentially establishing insightful conceptual connections between the special sciences.

Kemeny and Oppenheim formulate what has come to be known as a vocabulary count model of theoretical reduction. They adopt Thomas Nagel's terminology to formulate the basic principle of reduction in the sciences:

> In a reduction we are presented with two theories T_1 and T_2, and with the observational knowledge of today represented by the complex sentence

> [t] O...The theoretical vocabulary of T_2, Voc(T_2), contains terms which are not in Voc(T_1)...But it turns out that T_1 can explain all that T_2 can, and it is no more complex. Hence we drop T_2 from our body of theories, and strike out all the terms in Voc(T_2) which are not in T_1. Then we say that T_2 has been reduced to T_1.[1]

The authors consider four definitions to bring precision and clarity to the concept of theoretial reduction in the form of a rational reconstruction of the general requirements for an adequate reduction.[2] They outline three conditions that are supposed to be sufficient to effect a reduction in the sciences from a secondary scientific theory T_2 to a primary scientific theory T_1. According to their interpretation, T_2 has been reduced to T_1 when:

(i) T_1 can explain all that T_2 can.

(ii) T_1 is no more complex than T_2.

(iii) Hence: Drop T_2 from our body of scientific theories, and strike out all the terms in Voc(T_2) that do not occur in Voc(T_1).[3]

The account has the form of an enthymematic practical syllogism, or a sequential procedure to follow in effecting a theoretial reduction. It considers the theoretical terms of any pair of theories under consideration, with reference to potential differences in the cardinalities of the sets of theoretical terms contained in competing reduction candidates. The theory is judged ontically most economical among those with equivalent explanatory competence and systematization, whose complete set of theoretical terms has the least cardinality. We assume whatever systematization Kemeny and Oppenheim expect in their requirement (3) for a reducing theoretical explanation relative to any theory it reduces. When conditions (i) and (ii) are satisfied by theories T_1 and T_2, then we are instructed to implement directive (iii), by which the theoretical reduction of T_2 to T_1 is supposed to be achieved.[4]

---

[1] Kemeny and Oppenheim 1956, 9.

[2] Kemeny and Oppenheim 1956, 7: "As this process has been the subject of much philosophical controversy, it is the task of the philosopher of science to give a rational reconstruction of the essential features of reduction." See Swanson 1962, Schaffner 1967, and the papers collected in Agazzi 1991 and in Milkov and Peckhaus 2013.

[3] Kemeny and Oppenheim 1956, 8-9.

[4] Nagel (1951; 1961) is often acknowledged as the chief exponent of classical reduction in the sciences. In 1951, 299, Nagel distinguishes between reduction principles of definability and derivability. Feyerabend (1962) attacks the principle of derivability in Nagel's analysis of scientific reduction on the grounds that the meaning-invariance it presupposes does not obtain in possible instances of reduction. See also Coffa 1967, 500. Kemeny and Oppenheim's

2. Resilience of the K-O Model

The continuing intrigue of the K-O model depends largely on the fact that it features what is arguably the only objective measure of anything belonging to reduced theories having been literally *reduced* conceptually or in the cardinality of a new science's referential domain of existent objects. The K-O model proposes to count the theoretical term syntax tokens in the complete statements of a theory's putative propositional truths, and in particular to tally up the theoretical term tokens in the propositions that each theory advances as true explanations. When the relevant propositions are written out, assuming we know which terms are theoretical, we can simply highlight every occurrence of a theoretical term syntax token as though the words were presented in a two-dimensional matrix.

The > -1 algebra for K-O reductions applied to theoretical term tokens in a random theory's inscribed explanations nevertheless reveals the limitations of a purely syntactical albeit the only objective criterion of theoretical reduction. The implication is that the K-O model, whatever its fate at the hands of previous lines of criticism, and regardless of its current reputation and range of philosophical acceptance and acknowledged application, or the reverse, should either be: (a) rejected and replaced by a metatheory that interprets theoretical reductions in terms of the *meanings* of theoretical terms in reduction candidate theories, and not just the syntax of the theoretical term tokens scattered among a theory's propositions; or else (b) a major overhaul of the K-O model would be needed to accommodate semantic as well as purely syntactical dimensions of theoretical reductions from one choice of theoretical explanatory propositions to another. If meaning in the relevant theoretical expressions cannot be understood as purely objective, then a further apparently inescapable implication is that theoretical reduction in the sciences is also not a purely objective relation, phenomenon or occurrence.

However tempting it may be to turn away from the K-O model as old-fashioned or unsuited to a significant number of recognized theoretical reductions, to follow a trend of disregard for its usefulness in contemporary philosophy of science, to the point where few have studied its details, the K-O account of theoretical reduction cannot be so easily discounted, even as it braves indifference. The model succeeds in its most general form despite criticism and neglect because in the end it interprets theoretical reduction as involving a literal comparative numerical reduction in the theoretical vocabularies of competing scientific theories as the only objective measure of their comparative conceptual and consequently respective explanatory economies. To know to what concepts and entities a theory makes explanatory ontological commitments, the K-O model says that we must count the words that appear as

---

discussion ignores the principle of derivability for the most part, dealing with the principle of definability as bypassing the problem of meaning-invariance.

specific syntax items in a typically inscribed statement of the theory's explanations.

What else are we supposed to be able to do, if we are proceeding objectively, scientifically, in arriving at these metatheorical comparisons in support of the conclusion that one theory is reducible or has in fact been reduced to another? Even in the case of still living theorists who can further explicate their explanations by offering forth still more words to digest, and certainly with respect to the documented written heritage of theory development in a culture, there seems to be no available method except to read or otherwise process and evaluate the syntax in which a theory's explanations are expressed. Such considerations provide strong if not finally decisive justification for some form of the K-O model in the metatheory of theoretical reduction and its expected scientific methodology.

Elementary algebraic relations of > -1 govern the relative numbers of theoretical term tokens that belong to a theory than to the theory to which it is K-O reduced. As always, in the original K-O model, explanatory adequacy on both sides of theoretical reduction is presupposed, along with other condition to be met, so that theoretical reduction, as Kemeny and Oppenheim insist, can contribute to scientific progress. Differences in syntax token numbers in different choices of theoretical explanations can be understood as signifying both comparative economic differences in the numbers of concepts and entities to which a theory is ontologically committed, and, secondly, also, the comparative simplicity or complexity of such explanations, as reflected in the number of times a theoretical term must be employed within a theory's explanations in order to express its explanations.

3. Critique of the Comparative Vocabulary Reduction Model

An objection to the K-O model is that all three of the conditions in (i)-(iii) can be fulfilled in circumstances in which a theoretical reduction of scientific or other explanatory theory T_1 from T_2 in Kemeny and Oppenheim's sense is not effected.

What the underlying algebraic structure of the K-O model seems to reveal, demonstrated in a highly simplified application that nevertheless meets the K-O conditions, is that the K-O model is woefully inadequate in its inability to support correct evaluations of reduction relations in the overwhelming number of possible reduction candidates among choices of theoretical terms in the vocabularies of competing explanatory theories. The K-O model fails in particular for the vast number of random combinatorially available *syntactically token replicative* cases. The argument suggests that the K-O model, on these specific grounds, must either be rejected as an inappropriate interpretation of the comparative vocabulary concept of theoretical reduction, or, if the interpretation is judged correctly to capture the comparative vocabulary concept, then the idea of theoretical reduction itself must be rethought as a descriptive model of or prescriptive guideline for ideal scientific practice.

Theoretical reduction on the K-O model, as previously mentioned, is supposed to be progressive, resulting in theoretical explanatory improvements.[5] A genuine theoretical reduction must entail no loss in ability to explain phenomena when one theory is reduced to another, and the theory to which another is reduced must constitute a simpler or more economical way of explaining the same phenomena as the theory from which it is reduced. For Kemeny and Oppenheim, the simplification that is expected to result from a scientific theoretical reduction produces a greater economy in the number of terms in the scientific vocabulary. They begin by asking:

> What are the special features of reduction? Since it is to be progress in science, we must certainly require that the new theory should fulfill the role of the old one, i.e., that it can explain (or predict) all those facts that the old theory could handle. Secondly, we do not recognize the replacement of one theory by another as progress recognize the replacement of one theory by another as progress unless the new theory compares favorably with the old one in a feature that we can very roughly describe as its simplicity...And the special feature of reduction is that it accomplishes all this and at the same time allows us to effect an economy in the theoretical vocabulary of science.[6]

The objection to this reasonable proposal is that the K-O conditions (i)-(iii) do not necessarily guarantee reduction in the sciences in the relevant sense of 'simplicity', by effecting a theoretical economy in the scientific vocabulary. The argument to demonstrate the limitations of the K-O model of reduction begins with an elementary secondary science T_2 in which the following conditions obtain between scientific principles (A,B,C,D) and theoretical observations (O_1,O_2,O_3,O_4):

(1) A explains O_1

(2) B explains O_2

(3) C explains O_3

(4) D explains O_4

Here there are four explanatory scientific laws in one-one correspondence with four observations to be explained. This is already an unrealistic simplification, because scientific laws are ordinarily assigned the task of explaining many observations, and several laws are often needed to explain a single observation. Needless to say, besides, most scientific theories additionally include more than four scientific laws. Although the theory is simplified in at least these ways, it should nevertheless serve the purpose of illustrating a

[5] Kemeny and Oppenheim 1956, 6: "The label 'reduction' has been applied to a certain type of progress in science."

[6] *Ibid*.

general point about the limitations of the K-O model of theoretical reduction. There can, after all, be theories as basic as this interpretation of T_2, and if the K-O reduction model does not work even in this simple case, then it should probably not be expected to provide correct results when extended to increasingly more complex and to that extent potentially more realistic applications.

Consider what the K-O model would call the theoretical 'terms' contained within or by means of which the four scientific laws (A,B,C,D) of T_2 are expressed.[7] Again, somewhat artificially for the sake of argument, suppose that the vocabulary of theory T_2, Voc(T_2), consists of the follwing *vocabulary matrix* of 'terms':

Voc(T_2) Secondary Theory (Nonreplicative Case)

A = {a,a',a'',a'''} explains O_1

B = {b,b',b'',b'''} explains O_2

C = {c,c',c'',c'''} explains O_3

D = {d,d',d'',d'''} explains O_4

Grammatically and in other ways formally well-formed combinations of these theoretical terms associated with each law make it possible to explain each correlated observation. The terms in A, for example, {a,a',a'',a'''}, are used to explain O_1, and so on for O_2, O_3, and O_4. Collectively, the terms belonging to the four laws are the theoretical vocabulary of T_2, Voc(T_2), and presented above in a matrix array.

4. Theoretical Terms Nonreplicative Cases

The question is how a reduction of a secondary theory T_2 to a primary theory T_1 can be effected according to the K-O comparative vocabulary model. Two patterns of reduction are distinguished, designated as 'replicative' and 'nonreplicative.' As a paradigm of the nonreplicative case, to begin explaining the difference, suppose that T_2 above is reduced to T_1, where T_1 consists of the scientific laws (E,B,C,D), and where law E does a more economical job of explaining observation O_1, by virtue of containing only three terms (e,e',e''). The theoretical vocabulary of T_1 is thus:

Voc(T_1) Primary Theory (Nonreplicative Case)

E = {e,e',e''} explains O_1

B = {b,b',b'',b'''} explains O_2

[7] Observational or theoretical or both or neither; 'terms' simpliciter. See Jacquette 2004.

C = {c,c',c'',c'''} explains O_3

D = {d,d',d'',d'''} explains O_4

It appears that:

(i) T_1 by hypothesis can explain all that T_2 can in explaining O_1-O_4.

(ii) T_1 is no more complex than T_2, for it contains the same number of laws, and it contains fewer scientific terms ($Voc(T_1) = x = 16$, $Voc(T_2) = x - 1 = 15$).

(iii) Hence, we can drop T_2 from our body of theories, replace it with T_1, and strike out all the terms in the relevant vocabulary that occur in $Voc(T_2)$ but that do not occur in $Voc(T_1)$, viz.: the terms (a,a',a'',a'''). In this way we eliminate all of law A which has become superfluous in explaining O_1 after the discovery, verification, or acceptance of the more economical law E, by eliminating its theoretical terms.[8]

The fulfillment of these three conditions of the comparative vocabulary K-O model qualifies this first example as a genuine instance of theoretical reduction on the authors' terms. For we have eliminated four terms from the theoretical vocabulary (a,a',a'',a'''), and added only three (e,e',e''). Conditions (i)-(iii) are satisfied, and the replacement of T_2 by T_1 represents a simplification and greater economy of theoretical vocabulary. The example is unproblematic in the sense that it effects what Kemeny and Oppenheim would regard as scientific progress in theoretical reduction. It is a nonreplicative reduction, by virtue of the fact that it does not involve the replication of distinct tokens of any single syntactical term type, distributed over the theory's explanatory propositions. We assume throughout in what follows that the comparative vocabulary K-O model preconditions of explanatorily covering all the relevant observational data and being at least as well systematized (whatever this is finally understood to mean) when a reducing theory replaced a reduced theory.

5. Theoretical Term Tokens in Vertically Replicative Cases

The nonreplicative case is well-behaved but statistically atypical of scientific reductions. The percentage of possible nonreplicative reductions, supporting a matrix of four laws consisting of four scientific terms each, is swamped by the percentage of possible replicative cases in which conditions (i)-(iii) are satisfied,

[8] Kemeny and Oppenheim are hesitant about utterly eliminating superfluous terms from Voc(T2). See 1956, 17, note 3. Assume that the terms are definitely proven superfluous, thereby avoiding extralogical questions of convenience in holding on to technically unnecessary theoretical terms. We might consider, however, that such terms not be entirely eliminated from broader vocabularies of terms useful in strictly nonscientific explanations (e.g., in lay or historical explanations of scientific theories).

but in which nothing that Kemeny and Oppenheim would allow intuitively to count as a genuine reduction results. The problem is illustrated by the following replicative application of the K-O theoretical reduction model.

Suppose for simplicity sake that primary science T_1 consists of the scientific laws (E,B,C,D), and that, as above, the scientific terms of law E = (e,e',e''). Then, where $Voc(T_2) = x$, $Voc(T_1) = x - 1$ theoretical terms belonging to the two vocabularies. Now suppose also that the matrix of theoretical terms for both T_2 and T_1 for laws (B,C,D) in the nonreplicative case is not identical to the matrix of theoretical terms in the following replicative case. We stipulate again that E explains O_1, B explains O_2, C explains O_3, and D explains O_4, when T_2 is reduced to T_1. We permit restricted replication of scientific terms vertically in the matrix, but do not consider horizontal replication. To be precise, we specify the scientific terms of the two modified theories in the replicative case in this way:

$Voc(T_2^r)$ Secondary Theory (Replicative Case)

A = {a,a',a'',a'''} explains O_1

B = {b,b',b'',b'''} explains O_2

C = {c,a',c'',c'''} explains O_3

D = {d,d',d'',d'''} explains O_4

The theory is replicative in an obvious sense, because laws A and C share a single term a', rather than each containing completely different distinct scientific terms. The secondary theory is now K-O theoretically reduced to:

$Voc(T_1^r)$ Secondary Theory (Replicative Case)

E = {e,e',e''} explains O_1

B = {b,b',b'',b'''} explains O_2

C = {c,a',c'',c'''} explains O_3

D = {d,d',d'',d'''} explains O_4

There are difficulties for the K-O model that the replicative case immediately brings to light. Condition (i) is satisfied because both T_2 and T_1 adequately explain O_1-O_4. Condition (ii) is also satisfied because $Voc(T_2) = x = 16$, $Voc(T_1) = x-1 = 15$ scientific terms. When we attempt to fulfill condition (iii), however, as the K-O model requires whenever conditions (i) and (ii) are met, an interesting problem arises.

Satisfying condition (iii), we drop T_2 from our body of theories, and strike out the now superfluous theoretical terms in Voc(T_2) that do not appear in Voc(T_1). In the nonreplicative case we eliminate four terms (a,a',a'',a''') and add only three terms (e,e',e''), so that fulfillment of conditions (i)-(iii) effect a genuine reduction in such a way as to represent scientific theoretical progress. By contrast, in the replicative case, if we strike out the scientific terms that occur in Voc(T_2) that do not occur in Voc(T_1), we can strike out only the terms (a,a'',a'''), but not (a'), because (a') also occurs as a restricted vertical replication instance in Voc(T_1). After all, condition (iii) instructs us only to strike those terms from Voc(T_2) that do *not* occur in Voc(T_1). Law C in T_2 and in T_1 here consists of the terms (e,a',e'',e'''). Thus, we can only eliminate three terms (a,a'',a''') from the scientific vocabulary of T_1 in Voc(T_1).

However, since we have also added three terms to Voc(T_1) (e,e',e''), then, despite satisfying K-O model conditions (i)-(iii), no real *reduction* has been effected in the relevant simplification (comparative economy) sense of the K-O model, because the theoretical vocabulary of T_1 has not been simplified or made more economical than that of T_2. The net economy of a K-O comparative theoretical vocabulary reduction is necessary for the kind of progress that is supposed to characterize a genuine reduction through the replacement of one scientific theory by another. Kemeny and Oppenheim are quoted above as insisting:

> ...we do not recognize the replacement of one theory by another as progress unless the new theory compares favorably with the old one in a feature that we can very roughly describe as its simplicity... And the special feature of reduction is that it accomplishes all this and at the same time allows us to effect an economy in the theoretical vocabulary of science.[9]

Thus, there are instances in which all three conditions of the K-O model of theoretical reduction are satisfied, but where the theory that follows upon fulfilment of the conditions does not constitute a genuine theoretical reduction, given all that Kemeny and Oppenheim have informally to say about the requirements. The reason is that no economy in the scientific vocabulary and therefore no progress in science results when the conditions are satisfied in some term replicative applications. Is the replicative case significant? Can we ignore the problems it poses in light of the usefulness of the nonreplicative cases and the limited possibilites of the replicative case, the unlikelihood that it will appear among the reductions of otherwise methodologically scrupulous systematized theoretical explanations? It is easy to see that the replicative case is not a degenerate construction, because the percentage of its occurrences in a body of scientific theories projecting a matrix of theoretical terms as they appear

[9] *Ibid*., 6.

in two dimensions, containing theoretical terms in the relevant propositions both horizontally and vertically in a list of the theory's putative truths, is enormous compared to the alternative. The replicative case vastly outnumbers the nonreplicative cases of theoretical reduction on the K-O model. Moreover, important actual T_2 secondary theories in the history of science and in contemporary theoretical explanation almost always (98-99% of the logically possible cases) embed a disqualifying vertical syntactical replication of theoretical term tokens.

6. Algebraic Parameters of K-O Theoretical Reduction

More definite mathematical significance can be offered in support of this criticism of the K-O model of comparative theoretical vocabulary reduction, by comparing the percentage of possible nonreplicative cases against the percentage of possible replicative cases, using the same simplified matrix of sixteen scientific terms assigned in sets of four each to each of four scientific laws.

Suppose that the class of nonreplicative cases and the class of replicative cases logically exhaust the total possible instances of theories potentially entering into a theoretical reduction relation. In the nonreplicative case, there are in the simplified case precisely 1,820 possible combinations of terms available for nonreplicative reductions satisfying K-O model requirements (i)-(iii). This is determined combinatorially by the equation $n!/r!(n–r)!$, relying on the same pool of scientific terms, where n = the number of terms in the matrix, and r = the groupings of those terms for each law or horizontal coordinate of the matrix. In our simplified model, n = 16 and r = 4. Some of these possible configurations of scientific terms are uninteresting, such as the difference between (a,a',a'',a''') and (a,a'',a''',a'). So the importance of this mathematical information is not found in the absolute value of the cardinality of possible configurations, but in the ratio obtained by comparison of this indicated number with the total number of possible configurations permitted by the conditions of restricted vertical syntactical replications of term tokens in the replicative case.[10]

In the replicative case, further simplified to permit vertical but not horizontal replication in the matrix of scientific terms, there are 172,900 possible configurations of scientific terms. This is determined by the general formula $L(n!/(n–r)!)-(n!/r!(n–r)!)$, where, as before, n and r equal respectively the number of terms in the matrix and the size of the groupings of the subsets of

[10] Calculation of these values is extremely oversimplified in the application, which is atypical of theories and laws in propositional explanations actually used by scientists. Here only four laws are involved, with four terms each, and those laws stand in a one-one relation or correspondence with the observations they explain. The difficulties are compounded in cases dealing with more true-to-life scientific theories.

those terms, and where L is the number of scientific laws or horizontal coordinates of the matrix. The ratio of these two values of possible configurations in the matrix of sixteen scientific terms of the secondary theory to be reduced, where the nonreplicative case is compared to the total possible configurations in the universe of discourse, including both replicative and nonreplicative cases, is 0.0104166. This is to say, that in the limits of our universe of discourse of total possible nonreplicative and replicative cases, only 1.04166% can be of the sort we have called workable nonreplicative cases. The remaining 98-99% of such possible cases are unworkable because they are replicative. These are the sort of difficulties we encounter if we attempt to apply the K-O comparative vocabulary interpretation of theoretical reduction to cases of restricted vertical replication of scientific terms in a scientific theory's vocabulary matrix. The replicative instances are transparently generated by considering more than one, in fact by considering all the permitted replications in the matrix, in contrast with our simplified model in which only a' is replicated exactly once. It seems appropriate to conclude that the restricted utility (to 1.04166% of all possible instances) of the K-O comparative vocabulary model of theoretical reduction in the sciences warrants either its total rejection or major fundamental redesign.

7. Vertical Replication of Theoretical Terms in Actual Scientific Theories

We cannot prove, but we can suggest by way of examples selected entirely at random from ancient and contemporary scientific documents, that many if not most scientific theories contain laws that exemplify vertical replication of theoretical terms.[11] For these examples, and many like them, the K-O interpretation of theoretical reduction cannot be used to describe or guide a reduction to some primary science T_1.

First, Galileo writes in *De Muto Accelerato* (c. 1590):

1) …*bodies* of the same material but of unequal *volumes* move (in natural *motion*) with the same *speed*.

2) …when solids lighter than water are completely immersed in water, they are carried upward with a force measured by the difference between the weight of a *volume* of water equal to the volume of the submerged *body* and the *weight* of the *body* itself.

3) …if we wish to know at once the relative *speeds* of a given *body* in two different media, we take an amount of each medium equal to the

[11] Horizontal replications are also possible and frequently occur. For several reasons they are not considered in calculating the possible configurations of replicated terms in a matrix that specifies the scientific vocabulary of a theory Voc(Tn). The mathematical formula for computing the number of possible unrestricted replications of n terms is simply n! (n-factorial).

> *volume* of the *body* and subtract the *weight* (of such amounts) of each medium the *weight* of the *body*.
>
> 4) …in the (natural) downward *motion* of *bodies* the ratio of the *speeds* is not equal to the ratio of the *weights* of the *bodies*…[12]

While in a modern if not especially recent tried-and-true 1968 genetics textbook taken down from the shelf we find:

> 1) [*Chromosomes*] duplicate precisely and divide equally in mitosis, furnishing each cell with a full complement of *chromosomes*.
>
> 2) Their behavior in *meiosis* accords with our expectations of heredity — that it is due to contributions from both parents.
>
> 3) Their random mixing and crossing over during *meiosis* provides an important source for the observed variables between individuals.
>
> 4) In addition…*chromosome* abberations can be associated with the inheritance of specific characteristics. [13], [14]

Such instances are typical rather than exceptional among the groupings of propositions in an explanatory theory in which syntactically identical term tokens appear in several of the propositions according to the pattern we have referred to as vertical replication. There is usually a network of token syntactical linkages among the propositions advanced for purposes of theoretical explanation in a theory, reflected in the matrix of each specific theory's vocabulary of theoretical terms, targeted by the K-O theoretical reduction model for comparison in establishing theoretical reduction and reducibility relations between any two or more competing explanatory theories.

8. Countercritique of Theoretical Term Reduction Model Objections

Problems of several kinds might be raised against the matrix analysis of theoretical terms in a scientific theory. We conclude by addressing two such complaints, both of which seem dangerous, but on consideration neither of which seems to be especially compelling. The second criticism pinpoints exactly the philosophical difficulties that seem to be entailed whenever theoretical reduction procedures are described by ambiguous reference to syntactical 'terms,' as the main objection to the K-O comparative vocabulary model of theoretical reduction.

---

[12] Galileo 1960 [1590], 29, 33, 35.

[13] Strickberger 1968, 48.

[14] Consult the laws of a theory in almost any ancient or contemporary scientific text. For example, Galileo 1933, 203, 209, 218. Newton 1972 [1726]. Bent 1965, 15 (citing Joule's paper "On the Mechanical Equivalent of Heat" (communicated by Michael Faraday to the Royal Society in 1849)). Bloss 1971. See also Kimbrough 1979.

Criticism 1. It might be objected that counterexample replicative cases do not appropriately fit the K-O model of theoretical reduction, because if no genuine reduction is effected, that just means that condition (ii) is not satisfied. This reasoning misses the point of the argument. For initially condition (ii) is not and is not supposed or expected to be satisfied. $Voc(T_2) = x$ (16 terms), and $Voc(T_1) = x - 1$ (15 terms). The K-O interpretation then instructs us to enact condition (iii). However, when we carry out condition (iii), as previously observed, we do not effect a reduction of theories by simplification or economy in the theoretical vocabulary. Criticism 1 accordingly overlooks the precise way in which the K-O model authors have instructed us to use their schema as a kind of decision procedure, and we have followed these instructions in constructing our counterexample to criticize it on its own own erritory, and literally in its own terms. The simple-minded example presented above contains only one replication of a single term a' in the matrix of the relevant theoretical vocabulary. If two or more such terms, beginning with a' and b'', were to be vertically replicated within the matrix, then in a still significant percentage of total possible cases, condition (ii) would turn out to be initially satisfied, but no genuine reduction would result. T_1 would be 'more complex' than T_2, even under the deliberate misconstrual of the intentions of the K-O interpretation on which Criticism 1 is based.

Criticism 2. A more serious objection holds that we have no business counting a' as it occurs in law A (in T_2) as a denumerably distinct term from a' as it occurs in law C (also in T_2). If this is true, then the problems of the replicative case considered above disappear. For then $Voc(T_2) = 15$ instead of 16, and $Voc(T_1) = 15$ also, for the same reasons as before. Thus, if the argument is to get off the ground, we must eliminate a term in $Voc(T_1)$ in order to initially satisfy condition (ii).[15] Suppose that we respecify law E to consist of (e,e') instead of (e,e',e''), implying that $Voc(T_1) = x - 1$, once again, this time = 14. If we fulfill condition (iii) under these circumstances, then we do in effect what looks to be a genuine theoretical reduction according to the K-O model. We eliminate three

[15] Theory T_1 being 'no more complex than' T_2 may be sufficiently ambiguous to dismiss this objection. If theories are constructive propositional entities, then, if the propositions of one are simpler than those of the other, then the first theory should be simpler than the second. For propositions to be simpler than one another in scientific usage syntactically speaking covering the same explanatory obligations with competence and systematization can only be a matter of the number of countable theoretical terms contained in the reducing theory's explanatorily competent and well-systematized propositional replacement for the reduced theory's explanations. If you add more terms, you increase the grammatical combinatorics and hence the greater complexity of explanations. The greater number of theoretical term tokens, the greater potential for complexity, when its explanations stacked up against those of the reducing theory with a smaller syntactical theoretical term cardinality.

terms (a,a'',a''') from Voc(T_2) that do not occur in Voc(T_1), while adding only two terms (e,e'). In this application, we simplify and economize the theoretical vocabulary in replacing T_2 with T_1, thereby ridding the referential domain of as many corresponding theoretical concepts and objects. It may therefore seem that the application is an example of the nonreplicative case after all. Why indeed should we count two theoretical term tokens, a' and a' (say, in classical mechanics or kinematics, 'force' and 'force,' 'mass' and 'mass') as denumerably distinct terms? Why should we count these as two terms instead of two replications of a single term?

The best reason for considering these terms as distinct is that the K-O model of theoretical reduction in conditions (i)-(iii) and the authors' surrounding informal discussions, deal solely with 'terms' and not with the meanings the terms might be assigned. This means that tokens of the very same term vertically replicated with the matrix of a theoretical vocabulary can potentially mean radically different things. We ought for safety sake then to denumerate these replicated terms as discrete and distinct entities in determining the number of theoretical terms in a theory's vocabulary matrix. Naturally, it is only good theoretical practice not to use syntactically indistinguishable term tokens within a descriptive and explanatory vocabulary as having different meanings. Unfortunately, there can be no logical guarantee that this is not the case with respect to any particular theory and its theoretical terms. Moreover, the K-O model, as we have seen, makes no provision for sanitizing the terms in the vocabulary of a theory in a theoretical reduction relation on semantic grounds on the basis of the meanings of replicated theoretical term tokens within the theoretical vocabulary matrix, prior to determining whether or not conditions (i)-(ii) are satisfied, and on the strength of meeting those requirements implementing condition (iii). Nor is this the problem of meaning-invariance that Paul Feyerabend raises, in which term types are thought to change meanings holistically when extended across different theoretical frameworks.[16] We refer only to term tokens composed of the identically same letters or symbols in the same order and their grammatical variants that are deliberately or even unnoticed assigned different meanings by default within a single theoretical framework. Under ideal circumstances, such ambiguities and equivocations could not arise; although in an ideal world theoretical reductions would be unnecessary anyway, since all theories would already be maximally reduced to the minimal necessary theoretical structures and the matrix of their univocally replicating theoretical terms.

An intuitively trivial example, that the K-O reduction model nevertheless does not exclude, if one can forgive the awful puns, projects a set of laws in a biological theory containing the terms 'mole' meaning 'a burrowing mammal,' 'an

[16] Feyerabend 1962, 34, 41-43.

epidermal growth of tissue,' and 'a unit of measure, especially volume.'[17] There can obviously be more subtle differences of meaning in what seem to be identical terms vertically and even horizontally replicated in any set of explanatory propositions. The metatheoretical choices here are few. The discussion has led to recognizing the following two outstanding alternatives. We can either: (a) Recalculate by stipulation supported in argument syntactically replicated term tokens as distinct entities in the vertical coordinates (and perhaps also in the horizontal coordinates) of matrices containing the vocabularies of theoretical terms belonging to specific scientific and other kinds of explanatory theories. This option has already revealed its limitations, for it is precisely the condition of the above replicative term counterexamples to the K-O model, and as such offers no respite from its damaging conclusions for the K-O model; or (b) Conclude that the K-O model be rejected outright and in its entirety, if it cannot be amended to deal adequately semantically somehow with the meanings of theoretical term tokens in a theory's theoretical vocabulary, and not just with the syntactical forms of symbols that collectively include all the tokens of the theory's theoretical terms in any single statement of the theory's totality of putative propositional truths or at least its fundamental principles or axioms, also characterizable as the theory's propositional or thetic substance or content.

What continues to fascinate about the K-O model of theoretical reduction is its confident assumption that the possibility of an episode of theoretical reduction in the history of science can only be objectively made in supposedly purely syntactical terms of competing theoretical vocabularies, in which one theory comes to be reduced by and to another. Naturally, it is the relative cardinalities of theoretical syntax items in a larger context of all theoretical explanations as they are affected by the inclusion of the reduced or reducing theory that matter. A reducing theory in genetic biochemistry might make use of many theoretical concepts that are already part of chemistry, and use overall more theoretical terms in its explanations of a predecessor pre-DNA biological theory of genes, but still result in an integrated scientific network of explanations in which overall the number of syntactically distinct theoretical terms is diminished. The applications we have considered must all be considered accordingly as miniaturized versions of the complete scientific explanatory situation before and after a reduction, in which the total number of theoretical terms are compared when a theoretical reduction is considered. They are on each side the before and after theoretical term portraits of the reduced and reducing theories in the broadest context representing all theoretical terms in all theoretical explanations.[18]

[17] Or consider the less trivial fact that 'gram' in a chemical theory can mean either 'weight' or 'mass,' or, if the term is replicated, might mean both in different laws.

[18] I am grateful to several anonymous readers who have offered useful suggestions for improvement of previous drafts of the essay.

References:

Agazzi, Evandro, ed. 1991. *The Problem of Reductionism in Science*. Colloquium of the Swiss Society of Logic and Philosophy of Science. Zürich 18-19 May 1990. Dordrecht: Springer (Kluwer Academic Publishers).

Bent, Henry A. 1965. *The Second Law: An Introduction to Classical and Statistical Thermodynamics*. New York: Oxford University Press.

Bloss, F. Donald. 1971. *Crystallography and Crystal Chemistry: An Introduction*. New York: Holt, Rinehart, and Winston, Inc.

Coffa, José Alberto. 1967. "Feyerabend on Explanation and Reduction." *The Journal of Philosophy* 64: 500-508.

Feyerabend, Paul K. 1962. "Explanation, Reduction, and Empiricism." In *Minnesota Studies in the Philosophy of Science*, 3, 28-97. Minneapolis: University of Minnesota Press.

Galilei, Galileo. 1933. *Dialogues Concerning Two New Sciences*. New York: The Macmillan Company.

____. 1960 [1590]. *De Muto Accelerato*. Madison: University of Wisconsin Press.

Jacquette, Dale. 2004. "Theory and Observation in the Philosophy of Science." *Logical Analysis and History of Philosophy* 7, special issue on History of the Philosophy of Nature ('Geschichte der Naturphilosophie'): 177-196.

Kemeny, John and Paul Oppenheim. 1956. "On Reduction." *Philosophical Studies* 7: 6-19.

Kimbrough, Steven Orla. 1979. "On the Reduction of Genetics to Molecular Biology." *Philosophy of Science* 46: 389-406.

Milkov, Nikolay and Volker Peckhaus, eds. 2013. *The Berlin Group and the Philosophy of Logical Empiricism*. Boston Studies in the Philosophy and History of Science. Berlin: Springer Verlag.

Nagel, Ernest. 1951. "Mechanistic Explanation and Organismic Biology." *Philosophy and Phenomenological Research* 11: 327-338.

____. 1961. *The Structure of Science*. London: Routledge & Kegan Paul Ltd.

Newton, Isaac. 1972 [1726]. *Isaac Newton's Philosophiae Naturalis Principia Mathematica*. The Third Edition (1726) with Variant Readings. 2 Volumes. Assembled and Edited by Alexandre Koyré and I. Bernard Cohen with the assistance of Anne Whitman. Cambridge: Harvard University Press.

Schaffner, Kenneth F. 1967. "Approaches to Reduction." *Philosophy of Science* 34: 137-147.

Strickberger, Monroe W. 1968. *Genetics*. New York: The Macmillan Company.

Swanson, J.W. 1962. "On the Kemeny-Oppenheim Treatment of Reduction." *Philosophical Studies* 13: 94-96.

Revelation and the Essentiality of Essence

Franck Lihoreau

Abstract: It is usually agreed that the *Revelation Thesis* about experience – the idea that the knowledge we gain by having an experience somehow "reveals" the essence, or nature, of this experience – only requires that we know the essence of the experience, not that we know, of this essence, that it is the essence of the experience. I contest this agreement. In the light of what I call the "Essentiality of Essence Principle" – the principle that whatever is in the essence of something is also essentially so – I argue that the Revelation Thesis does require that we know, of the essence of an experience, that it is the essence of the experience, and draw some conclusions about the plausibility of that thesis.

Keywords: essentiality of essence principle, experience, qualia, revelation thesis

1. The Revelation Thesis about Experience

This paper is concerned with the so-called 'Revelation Thesis' about experience,[1] the idea that the knowledge we gain by having an experience 'reveals' the essence, or nature, of this experience, what it is or what it takes to be that experience.

Often formulated in relation to *qualia* – the phenomenal properties instantiated in our experiences – as the claim that *when we have an experience with quale* Q*, we thereby (are in a position to) know the essence of* Q,[2] the relatively few philosophers who have paid attention to the thesis have disagreed on a number of points, the most important of which including the questions whether *Revelation* is to be considered part of our folk-psychological notion of experience (Lewis 1995 thinks so; Stoljar 2009 thinks not), and whether it should be dismissed as inconsistent with physicalism (Lewis 1995 and Stoljar 2009 think so; Damnjanovic thinks not). (I shall briefly return to these questions in the last section.)

Yet, those philosophers all seem to agree on one point: on a proper formulation of the Revelation Thesis, the expression 'knowing the essence, *F*, of a

[1] So named after Johnston 1992, also referred to, by Lewis 1995, as the "Identification Thesis".
[2] Following Williamson 2000, Damjanovic (2012, 71-72) suggests interpreting 'being in a position to know' the essence of *Q* as meaning that the essence of *Q* is 'open to one's view,' so that 'all that must be done to know that [essence]' is for us to reflect on our experiences. For our limited purposes, an intuitive reading of the expression will work as well.

© Symposion, 1, 1 (2014): 69-75

quale *Q*' is to be understood as 'knowing that *Q* is *F*,' and *not* as 'knowing that *F* is the essence of *Q*.' In other words, they agree that there is a substantial distinction to be made between a 'weak' (or *de re*) version and a 'strong' (or *de dicto*) version of the Revelation Thesis:

> (*Weak Revelation*) When we have an experience with quale *Q*, if it is in the essence of *Q* that *p*, we thereby (are in a position to) know *that p*.
>
> (*Strong Revelation*) When we have an experience with quale *Q*, if it is in the essence of *Q* that *p,* we thereby (are in a position to) know *that it is in the essence of Q that p*.[3]

and that by *Revelation*, we ought to refer to the former version, not the latter.[4]

It is the legitimacy of this agreement which I wish here to consider. And I shall contest it on the grounds that *Weak Revelation* entails *Strong Revelation* as a special case. So that if it is in the essence of quale *Q* that *P*, *Revelation* does not only require that the experiencer know that *P*, but also that he know that it is in the essence of *Q* that *P*. Or so shall I argue in the light of a principle not generally discussed, but which I take to be as natural as can be when it comes to reasoning about essences (or natures).

[3] In these two formulations and throughout the paper, *p* stands for any sort of proposition that can be true of a quale. I certainly acknowledge that distinguishing between knowledge of essential truths and knowledge of essential properties might be relevant in the context of a discussion of Revelation – Damnjanovic 2012 is a good illustration –, yet not with respect to the purpose of the present paper in which *p* will be used to refer to any kind of claim that can be true of a quale (property ascriptions, identity claims, etc.).

[4] Daniel Stoljar (2009), for instance, warns us that:

> Knowing the essence of experience might be interpreted as knowing (e.g.) *that this itch is F*, where *F* is *in fact* the essence of the experience; or it might be interpreted as knowing *that F is the essence of this itch* (Stoljar 2009, 119)

and makes it clear that we ought to adopt the first formulation (for reasons that will be addressed in the last section); followed in this by Nic Damnjanovic (2012), who explicitly assumes from the start that:

> If it is in the nature of the taste of peaches that *p* then *Revelation* only requires that we know that *p*, and not that it is in the nature of the taste of peaches that *p* (Damnjanovic 2012, 71, fn. 4).

On this point, most people who have lent some thought to the thesis have agreed with them. (An exception is Byrne and Hilbert 2007 who defend a form of *Strong Revelation*, but in relation to colours, rather than to experience. For this reason, this exception shall not detain us here).

2. The Essentiality of Essence

The principle I have in mind states that whatever is essential to something is also essentially so. To be precise, this 'Essentiality of Essence Principle,' as I shall call it, states that:

> (*Essentiality of Essence*) For any thing (or class of things) *X* and any proposition *p*, if it is essential to *X* that *p* then it is essential to *X* that it be essential to *X* that *p*.

(Remark. I will use the expressions 'it is essential to *X* that *p*' and 'it is in the essence of *X* that *p*' interchangeably.) This principle follows, it will be seen, from basic, uncontroversial considerations regarding the connections between essentiality and necessity.

The question whether the essential *versus* accidental property distinction can be properly rephrased in modal terms has been, and still is a matter of dispute – most notably between those who think that being essential basically amounts to being necessary (a view often associated with Kripke 1980), and those who contend that essential properties cannot be analyzed in modal terms but need be characterized otherwise (for instance in definitional terms as in Fine 1994, or in explanatory terms as in Gorman 2005).[5] There are, however, two undisputed and, I believe, undisputable claims that can be made, connecting essentiality with necessity on the one hand, accidentality with possibility on the other hand:

> ($E/\Box$) For any proposition *p* and any object (or class of objects) *X*, if if it essential to *X* that *p* then it is (metaphysically) necessary that *p*.
>
> ($A/\Diamond$) For any proposition *p* and any object (or class of objects) *X*, if if it accidental to *X* that *p* then it is (metaphysically) possible that not-*p*.

And these two claims, together with the following (also natural) claim:

> ($\neg E/A$) For any proposition *p* and any object (or class of objects) *X*, if *p* is true of *X* and yet it is not essential to *X* that *p*, then it is accidental to *X* that *p*.

are all that we need in order to derive the above Essentiality of Essence Principle.

To see this, let *A* be an object, *P* a proposition, and assume that *P* is not only true of *A*, but is also an essential truth about *A* (for instance, *A* could be water and *P* the proposition that water is made of H2O molecules). In other words, assume that:

[5] This is an oversimplified view of the debates. For a survey of the literature on the analysis of the essential/accidental property distinction, see Robertson and Atkins 2013.

(1) It is essential to *A* that *P*

Then, by ($E/\Box$):

(2) It is also necessary that *P*.

On the other hand, assume, for *reductio ad absurdum*, that:

(3) It is not essential to *A* that it is essential to *A* that *P*.

Then, by ($\neg E/A$):

(4) It is only accidental to *A* that it is essential to *A* that *P*.

This, by ($A/\Diamond$), implies that:

(5) It is possible that it is not essential to *A* that *P*.

Which, by (A/◊) again, implies that:

(6) It is possible that it is possible that not-*P*.

This, of course, is incompatible with (2) as a matter of basic modal logic. By *reductio*, (3) is false and as a consequence:

(7) If it is essential to *A* that *P,* it must also be essential to *A* that it be essential to *A* that *P.*

(So, for instance, if it is essential to water that it be made of H2O molecules, then it is also essential to water that water be essentially made of H2O molecules.)

Since *A* and *P* are arbitrary in the foregoing reasoning, we can thereby see that the Essentiality of Essence Principle follows from fairly uncontroversial claims regarding essentiality, accidentality and modality.[6] It is this principle that will play the pivotal role in my argument against considering *Weak Revelation* and *Strong Revelation* as if they were two substantially different theses.

3. From Weak to Strong Revelation

In the light of the Essentiality of Essence Principle, it should clearly appear that *Weak Revelation* entails *Strong Revelation* as a special case.

For consider *Weak Revelation* again:

(*Weak Revelation*) When we have an experience with quale *Q* and it is in the essence of *Q* that *p*, we thereby (are in a position to) know that *p*.

Suppose that, in fact:

[6] Although I have derived the Essentiality of Essence Principle from claims connecting essentiality/accidentality talk with necessity/possibility talk, it is worth mentioning that this principle is also a theorem of the 'logic of essence' associated with Kit Fine's non-modal, definitional approach to the concept of essence, as found in Fine 1995 and Correia 2000. I take this as a further reason to endorse the principle.

(1) We have an experience with quale *Q*

and assume that:

(2) It is in the essence of *Q* that *P*

for some arbitrary *Q* and *P*.

From this latter assumption, (2), and the Essentiality of Essence Principle, it follows that:

(3) It is in the essence of *Q* that it is in the essence of *Q* that *P*.

Now, given (3), we can substitute 'it is in the essence of *Q* that *P*' for *p* in *Weak Revelation* above, so that we get:

(4) If we have an experience with quale *Q* and it is in the essence of *Q* that it is in the essence of *Q* that *P*, we thereby (are in a position to) know that it is in the essence of *Q* that *P*.

And from this, together with the conjunction of (1) and (3), it follows by *modus ponens* that:

(5) We thereby (are in a position to) know that it is in the essence of *Q* that *P*.

We may therefore conclude that:

(6) If we have an experience with quale *Q* and it is in the essence of *Q* that *P*, we thereby (are in a position to) know that it is in the essence of *Q* that *P*.

That is, we get *Strong Revelation*.

Therefore, assuming the truth of the Essentiality of Essence Principle, *Strong Revelation* is but a special case of *Weak Revelation*. In other words, if it is in the essence of a phenomenal property *Q* that *P*, the Revelation Thesis requires not only that we know *P*, but requires that we know *P* to be essential to *Q* as well. Hence, the agreement found in the literature on the supposedly substantial distinction between a weak and a strong version of *Revelation* is ill-grounded.

4. Is (Strong) Revelation Plausible?

Now, what bearings does this have on the plausibility of the Revelation Thesis as part of a theory of experience? We just established that *Weak Revelation* entails *Strong Revelation* as a special case. This should sound like bad news for advocates of the Revelation Thesis. Here are two reasons why.

(A) A first point that might be pressed is that, as Stoljar (2009) points out, *Strong Revelation* amounts to requiring experiencers to antecedently possess some concept of essence, while "surely it is implausible that those who are itchy require the concept of essence." Stoljar might be right on this but what he says needs a little supplementation.

(i) If, as Stoljar thinks, knowing something to be the essence of a given quale requires possessing the concept of essence *in general* (rather than the concept of the essence of this or that quale in particular) – which seems to be Stoljar's intended interpretation –, then I agree with him that *Strong Revelation* can hardly be true. Whatever meaning 'possessing the concept of essence' might have – whether we understand it as being able to articulate (some of) its conditions of application**,** or simply as being able to exercise, even unreflectively, that concept with (some) success, or as merely requiring that one have some notion of what makes a thing the thing it is –, people (even professional metaphysicians) surely have had lots of experiences long before they possessed the concept of what makes a thing the thing it is.

(ii) If, on the other hand, knowing something to be the essence of a quale requires possessing the concept of the essence of this quale *in particular* (rather than the concept of essence in general), this is likely to tense up the ears of anyone who thinks that the notion of qualia is to play a role in a proper understanding of experience. And the reason for this is that *Revelation* could easily be used in a sceptical argument against qualia talk. The argument in question would go like this. If knowing something to be the essence of a quale requires possessing the concept of this quale, then (*Strong*) *Revelation* entails that (1) one cannot have an experience with a quale *Q* unless one first has some notion of what makes *Q* the thing it is; but (2) one cannot have some notion of what makes *Q* the thing it is unless one could and did attend to *Q* in the first place; but, of course, (3) one cannot attend to *Q* at all before one has an experience with *Q*; therefore, (4) one could never have an experience with *Q at all!* – a result that a qualia theorist would not be willing to accept.

(Remark. Although such considerations incline me to agree with Stolar that *Strong Revelation* is implausible, I disagree with him in that he takes that implausibility to be a reason why proponents of *Revelation* – of which, by the way, he is not – should retreat to its *Weak* version. For if *Strong Revelation* just is a special case of *Weak Revelation*, as I have argued, the implausibility of the former might so well be taken as a reason to dismiss the whole idea of *Revelation* in itself as implausible.)

(B) One might also press another point against proponents of the Revelation thesis, a point that echoes Lewis's objection to the thesis. Considering it in its *Weak* version, Lewis has famously argued that *Revelation*, which he takes to be part of the folk theory of experience, should be rejected as false if physicalism is true, that is, "if qualia are physical properties of experiences, and experiences in turn are physical events." For "then it is certain that we seldom, if ever, [know the nature of] the qualia of our experiences. Making discoveries in neurophysiology is not so easy!" (Lewis 1995, 328). Lewis concludes that our folk theory of qualia has to be revised.

I do not wish here to consider whether physicalism is true or false, or even whether it is compatible with *Revelation* or not (see Damnjanovic 2012 for a

compatibilist position on this latter question). Yet, I am confident that anyone sympathetic to Lewis's argument will find it at least as effective when directed at *Strong Revelation*. For if *Strong Revelation* were true, it would not only be easy to discover the neurophysiological correlates of the qualia of our experiences – e.g. that this pain is a C-fiber firing –, but also to discover ontological truths about these correlations – viz. that it is in the essence of this pain to be a C-fiber firing. This is likely to sound incongruous to many ears. What is more, note that this incongruity would hold even if physicalism were false: whatever properties, events, be they physical or not, whatever truths are part of the essence of an experience, (*Strong*) *Revelation* entails that these properties, events, truths, are revealed to us as such. But making discoveries in the metaphysics of mind, we might say, is not so easy either!

I have mentioned these two points – there may well be many more, which I leave it to others to press – just to give a feel of how problematic *Revelation* might turn out, once established the implication from its *Weak* to its *Strong* version. Establishing this implication was my main purpose here. If the essence of an experience is ever revealed to us, then it is revealed to us as such.

References

Byrne, Alex and David R. Hilbert. 2007. "Colour Primitivism." *Erkenntnis* 66: 73-105.

Correia, Fabrice. 2000. "Propositional Logic of Essence." *Journal of Philosophical Logic* 29: 295-313.

Damnjanovic, Nic. 2012. "Revelation and Physicalism." *Dialectica* 66: 69-91.

Fine, Kit. 1994. "Essence and Modality." *Philosophical Perspectives* 8: 1-16.

____. 1995. "The Logic of Essence." *Journal of Philosophical Logic* 24: 241-273.

Gorman, Michael. 2005. "The Essential and the Accidental." *Ratio* 18: 276-289.

Johnston, Mark. 1992. "How to Speak of the Colours." *Philosophical Studies* 68: 221-263.

Lewis, David. 1995. "Should a Materialist Believe in Qualia?" *Australasian Journal of Philosophy* 73: 140-144. Reprinted in his *Papers in Metaphysics and Epistemology*, 325-331. Cambridge: Cambridge University Press. All references are to the reprinted version.

Robertson, Teresa and Philip Atkins. 2013. "Essential vs. Accidental Properties." *Stanford Encyclopedia of Philosophy*. http://plato.stanford.edu/entries/essential-accidental/ (Retrieved on July 14, 2014).

Stoljar, Daniel. 2009. "The Argument from Revelation." In *Conceptual Analysis and Philosophical Naturalism*, eds. David Braddon-Mitchell and Robert Nola, 113-138. Cambridge: MIT Press.

Williamson, Timothy. 2000. *Knowledge and Its Limits.* Oxford: Oxford University Press.

Skeptical Thoughts Concerning Explanationism and Skepticism

Clayton Littlejohn

Abstract: According to the explanationist, we can rely on inference to best explanation to justifiably believe familiar skeptical hypotheses are false. On this view, commonsense beliefs about the existence and character of familiar, medium-sized dry goods provides the best explanation of our evidence and so justifies our belief that we're not brains-in-vats. This explanationist approach seems *prima facie* plausible until we press the explanationist to tell us what the data is that we're trying to explain by appeal to our beliefs about external objects and how we could have access to it.

Keywords: dogmatism, evidence, explanationism, inference to best explanation, knowledge, skepticism

Introduction

Please consider two hypotheses:

> BIV: Your body consists of nothing more than a brain that's currently housed in a vat of nutrients in a lab in Nevada. Scientists are conducting experiments on your brain to see whether it's possible to systematically deceive you about the nature of your surroundings and about the past. They cause you to undergo experiences that are indistinguishable from the experiences you're having right now and induce false memories to trick you into believing various things about the past.

> Hands: Your body consists of much more than just a brain. You have a pair of hands that you can see when held out in front of you.

You know that these hypotheses cannot both be true. It's quite natural to think that you know that the Hands-hypothesis is true. It's also quite natural to think that you couldn't know the Hands-hypothesis to be true unless you were in a position to know that the BIV-hypothesis was false. Because of this, it seems quite natural to think that you're in a position to know that the BIV-hypothesis is false. Let's suppose that's right. The question that concerns us here isn't *whether* you can know that you're not a BIV, but *how* you could know this on the assumption that you can.

© Symposion, 1, 1 (2014): 77-87

An increasingly popular explanation of how you could know that a skeptical hypothesis is false is the *explanationist* explanation.[1] The gist of the explanationist view is that we can justifiably believe and know that we're not BIVs because the best explanation of our evidence (or features of our evidence) is an explanation that appeals to commonsense beliefs, such as the belief that you have hands.[2] Assuming that we do indeed have hands and that we're not BIVs, it's an inference to best explanation that provides the justification we need to know that we're not BIVs.

Some critics of this approach will criticize the idea that inference to best explanation provides justification or knowledge-level justification for our beliefs. For the purposes of this paper, let's bracket general concerns people have raised about the epistemic power of inference to best explanation. The question we shall focus on here is whether the explanationist view provides a satisfactory account of how we can know that a skeptical hypothesis like the BIV-hypothesis is false.

To understand the explanationist view, it will be helpful to briefly discuss a familiar skeptical argument and some standard responses to it. Let's start with the argument from ignorance:

The Skeptical Argument from Ignorance

P1. You cannot know that you're not a BIV.

P2. If you cannot know that you're not a BIV, you cannot know that you have hands.

C. Thus, you cannot know that you have hands.

You could simply accept the argument's conclusion for the reasons (or 'reasons') given. Some do have the strong intuition that we cannot know that we're not BIVs, so the skeptical view has something to be said for it insofar as it vindicates an intuition some of us surely have. The downside of skepticism, of course, is skepticism. Many of us think that it's at least possible to know things about the external world. Accepting the argument's conclusion is taken by many of us to be a significant cost. If you want to resist the skeptical argument, you could deny the argument's second premise. You might say that there are counterexamples to the closure principle and say that the case we're dealing with is a perfectly good counterexample to the general principle. The downside of this response is that the closure principle strikes many of us as being well motivated. While this view might vindicate the intuition that we can know that we have hands, it does so in a very costly way. Let's suppose we want to retain closure and want a view that allows for the possibility of knowledge of the

[1] See BonJour 1999, Feldman 2001, Lipton 2004, Lycan 2002, McCain 2014, Moser 1988, and Vogel 1990 for sympathetic presentations of the explanationist responses to skepticism.

[2] For a helpful discussion of explanationism that focuses on attempts to justify inference to best explanation, see Beebe 2009.

external world. Setting contextualist views aside, let's consider the prospects of an explanationist view that denies the argument's first premise.

If you deny the first premise of the skeptical argument, you'll face the non-trivial challenge of explaining how it could be false. I think it's helpful here to remember Pryor's 2000 advice. We shouldn't simply assume that the argument's first premise is correct. We should press those who accept the argument's first premise for an explanation as to why we should do likewise. One way to try to motivate it is by appeal to three ideas. The first is that we couldn't know that we're not BIVs unless we were justified in believing that we're not BIVs.[3] The second is that we couldn't be justified in believing that we're not BIVs unless we had adequate experiential evidence for believing that we're not BIVs.[4] The third is that we cannot have adequate evidence for believing that we're not BIVs.[5]

The first two points seem relatively unproblematic, but consider the third. Why shouldn't we say that our evidence *is* adequate for believing that the BIV-hypothesis is false? Perhaps this is what we should say. If we do say this, we'll be pressed to explain how our evidence could be adequate for supporting this belief. As many epistemologists see it, we'd have the same evidence for our beliefs regardless of whether the Hands-hypothesis or the BIV-hypothesis is true.[6] Doesn't that show that our evidence couldn't be adequate for believing the Hands-hypothesis and rejecting the BIV-hypothesis?

Although there are epistemologists who will say that you couldn't have adequate evidence for believing that you're not a BIV unless you had evidence that wasn't available to a BIV, the explanationist view is supposed to explain how the evidence we have could be adequate for our beliefs *even if* it's just the evidence we'd have if we were BIVs.[7] As they see it, the Hands-hypothesis provides the best explanation as to why our evidence is the way it is. If so, this explains why we have adequate evidence for believing the Hands-hypothesis and there would be no principled objection to rejecting the first premise of the skeptical argument.

---

[3] This is controversial, but let's set this aside for the time being as I'm going to talk as if knowledge requires justification because it's convenient to do so. Nothing of importance here hangs on any controversial claims about the relationship between justification and knowledge. By speaking as if knowledge required knowledge-level justification, I can simplify the discussion and avoid having to talk about skeptical challenges that challenge different epistemic statuses.

[4] For our purposes evidence will be understood as experiential or memorial. Let's assume for the purposes of this discussion that we don't have apriori justification for believing the BIV-hypothesis is false.

[5] See Pritchard 2012 and Pryor 2000 for discussion of this defense of the argument's first premise.

[6] Cohen 1984 and Conee and Feldman 2004 defend the thesis that you and your deceived BIV counterpart share the same evidence.

[7] See Littlejohn 2012, McDowell 1998, and Williamson 2000.

To flesh out the details of the explanationist view, the explanationist needs to answer two questions. First, by what criteria should we judge candidate explanations? Second, what is the Hands-hypothesis supposed to explain? The explanationist can recommend evaluating candidate explanations in terms of virtues like simplicity (i.e., the number and kinds of entities posited, the internal simplicity of the explanatory framework), fecundity (i.e., suggests further hypotheses that would be explanatorily useful), neatness (i.e., leaves fewer unanswered questions than rivals), or plausibility in terms of how well it coheres with our other justified commitments.[8] As for the features of our evidence that the explanatory hypotheses purport to explain, the explanationist might take the Hands-hypothesis to figure in the best explanation of the following (alleged) features of our experiential evidence:

F1: The coherence and regularity of experiences in a single sense modality.

F2: The coherence and regularity of experiences across more than one sense modality.

F3: It typically appears to us that external objects exist when we expect them to and these objects typically appear to be the way we expect them to be.[9]

The explanationist thinks that to justifiably believe or know that the BIV-hypothesis is false, we have to know that our experiential evidence has some combination of these features and discern that the best explanation of some combination of (F1)-(F3) is best explained by appeal to commonsense beliefs such as the belief that you have hands.

Explanationism vs. Dogmatism

It wouldn't be surprising if most readers agreed that the Hands-hypothesis provides the best explains the relevant features of our evidence. It also wouldn't be surprising if most readers agreed that inference to best explanation is generally capable of conferring justification. The tricky question is whether we need to rely on inference to best explanation to justifiably reject the BIV-hypothesis. As the view is understood here, the explanationist doesn't simply assert that inference to best explanation is one may amongst many to acquire justification for rejecting the BIV-hypothesis. The explanationist also thinks that rival views cannot provide an adequate account of how it is that we can justifiably believe that we're not BIVs.

To evaluate the explanationist view, then, we should consider a rival approach. Consider Pryor's 2000 dogmatist view. Like the explanationist, Pryor's dogmatist wants to say that we can have experiential evidence that provides sufficient evidential support for believing that the BIV-hypothesis is mistaken.

[8] See Lipton 2004 and Lycan 2002.

[9] BonJour 2003, Feldman 2004, McCain 2014, and Vogel 1990 mention these kinds of features in their discussions of explanationism.

Like the explanationist, Pryor's dogmatist doesn't think that the adequacy of our evidence for believing that we're not BIVs depends upon whether we have evidence that our BIV-counterparts do not. The disagreement between the dogmatist and the explanationist has to do with the power of particular experiences to justify the belief that we're not BIVs. It's possible, Pryor thinks, for a particular, one-off experience to justify the belief that you have hands. Once that belief is justified, you can justifiably infer that you're not a BIV, provided that you know that the BIV-hypothesis is incompatible with the Hands-hypothesis. Thus, on Pryor's view, you don't have to appeal to features like (F1)-(F3) to justify the belief that you're not BIV. On his neo-Moorean view, you can bypass that and justifiably infer that you're not a BIV by reasoning as follows:

The Neo-Moorean Argument

NM1: I have hands.

NM2: If I have hands, I'm not a BIV.

NMC: I'm not a BIV.

The explanationist and the dogmatist can agree that your justification for believing NM2 is apriori. They can also agree that your justification for believing that NMC derives from experience. Their disagreement concerns the power of one-off experiences to provide justification for believing NMC via this sort of inference. As the dogmatist sees it, there's no principled reason to think that the one-off experience you have when you see your hands is incapable of justifying your belief in NM1. Once we've accepted that, the dogmatist will say that there's then no principled objection to the idea that that experience provides sufficient evidential support for believing NMC. It's this second point that's the primary point of contention in the debate between the dogmatist and the explanationist.

If the one-off experience that you have when you see your hands is the sort of thing that could be sufficient to justify the belief that you have hands, the dogmatist thinks that it should also be sufficient to justify your belief that you're not a BIV. The explanationist thinks that the one-off experience cannot provide adequate evidential support for rejecting the BIV-hypothesis, but this puts them in a tricky spot, doesn't it? Let's consider two sufficiency claims:

S1: The one-off experience you have when you see that you have hands can provide sufficient evidential support for believing that you have hands.

S2: The one-off experience you have when you see that you have hands can provide sufficient evidential support for believing that you're not a BIV.

The dogmatist can run the following argument against the explanationist view:

The First Anti-Explanationist Argument

FAE1: The one-off experience you have when you see that you have hands can provide sufficient evidential support for believing that you have hands regardless of whether your experiences have features (F1)-(F3).

> FAE2: If your one-off experience as of hands can provide you with sufficient evidence to justifiably believe that you have hands regardless of whether your experiences have features (F1)-(F3), your one-off experience as of hands can provide you with sufficient evidence to justifiably believe that you're not a BIV regardless of whether your experiences have features (F1)-(F3).
>
> FAEC: Your one-off experience as of hands can provide you with sufficient evidence to justifiably believe that you're not a BIV regardless of whether your experiences have features (F1)-(F3).

If this argument is sound, it would deliver a serious blow to the explanationist view as it would show that the ability to justifiably reject the BIV-hypothesis does not depend upon whether there's some feature of our experiential evidence that's best explained by appeal to commonsense beliefs like the belief in the Hands-hypothesis.

Although the explanationist isn't committed to closure, we're exploring the possibility that the explanationist can rebut the skeptical argument without denying closure. It seems that *if* the explanationist accepts closure, they should accept (FAE2). The idea here isn't that the closure principle commits you to saying that (S2) follows from (S1). There are coherent views that accept closure, accept (S1), but reject (S2). These are views on which the justification for rejecting the BIV-hypothesis doesn't derive from (solely) experience but must be in place for experience to justify believing that you have hands.[10] These views differ from the explanationist view on offer insofar as they deny that the justification for rejecting the BIV-hypothesis derives wholly from the justification provided by experience. The reason that the explanationist who accepts closure should accept (FAE2) is that if (F1)-(F3) are necessary for having justification to believe that the BIV-hypothesis is false, these are necessary for having justification to believe propositions that are known to entail that the BIV-hypothesis is false.

It looks like the explanationist will have to reject the argument's first premise, (FAE1). This doesn't come without intuitive costs. In terms of the phenomenology, it doesn't seem like your belief that you have hands is formed as the result of an inference that rests on multiple observations.[11] It seems rather odd to think that the objects that are sitting in plain view are *ens theoria*. Wouldn't it take a powerful argument to show that the justification for believing

[10] For discussion of such views, see Silins 2005 who accepts a liberal view on which (S1) is true that isn't committed to the dogmatist view on which both (S1) and (S2) are true. As he argues, it's possible for someone to accept Pryor's account of the justification of mundane beliefs like the belief that you have hands while accepting something like Wright's 2004 account of the justified rejection of skeptical hypotheses. For a critical discussion of Silins' proposal, see Kotzen 2012.

[11] It's interesting that Feldman 2001, for example, insists that our evidence for rejecting the BIV-hypothesis is indirect since he seems to agree with Pryor in his 2004 that it's possible for particular experiences to provide non-inferential justification for the belief that we have hands. There seems to be a tension here because Feldman also endorses closure.

in such things requires the kind of inferential work that's required for inferring that the existence of the thing provides the best explanation of our evidence?

We shouldn't expect the explanationist to accept that this is a decisive counterexample to their view. Neutral parties might be put off by the idea that the justification for believing that you have the hands you seem to see before you really derives from an inference to best explanation, but perhaps the explanationist would reasonably claim that they have a good reason for rejecting (S1). They could say that we ought to reject (S1) because *if* we accept (S1) and accept closure, we'd either have to accept (S2) or we'd have to say that the justification for rejecting the BIV-hypothesis doesn't derive from our experiential evidence. They might admit that the rejection of (S1) is initially counterintuitive, but they might insist that it's a price worth paying in light of the fact that (S2) is so deeply counterintuitive. Is this a sufficient dialectical riposte?

If the explanationist is offering a principled response here, it seems that the principle that they're relying on couldn't be this one:

> ISE: Every justified belief is justified by virtue of an inference to best explanation.[12]

While this impossibly strong explanationist principle supports the idea that the belief in the Hands-hypothesis could only be justifiably accepted on the basis of an inference to best explanation, the impossibly strong explanationist principle is impossibly strong. At some point, mustn't we have a non-inferential way of knowing about our evidence and various features of our evidence so that we have a kind of direct access to something that we can raise questions about? Similarly, if the inference to best explanation that generates justification is supposed to explain features of our evidence, it seems that we would need to have a way to form justified beliefs about these features of the evidence and at some point it seems we'd need a non-inferential way of knowing about some of the features of our evidence.

The explanationist should appeal to a weaker principle, one that allows for the possibility of non-inferential knowledge and justification. The principle has to support the idea that the justification we have for believing the Hands- and BIV-hypotheses is generated by inference to best explanation but allow for the possibility of non-inferential justification for believing various things about our evidence. This would do the trick:

> ME: For any proposition that's incompatible with a skeptical hypothesis (e.g., the BIV-hypothesis) to be justifiably believed, it has to be justified by an inference to best explanation.

[12] An expalanationist view that incorporates ISE is even stronger than the view Lycan (2002, 417) dubs 'ferocious explanationism.' According to ferocious explanationism, the only ampliative inferences that generate justification derive their epistemic power from explanatory inference. This view allows for non-inferential justified judgments, but ISE appears to rule that out.

This modest explanationist principle supports the idea that you cannot justifiably believe that you have hands or that you're not a BIV unless the Hands-hypothesis (or something like it) is the something you see is part of the best explanation of your evidence. Unlike the impossibly strong explanationist principle, the modest principle doesn't tell us that you need to rely on inference to best explanation to justifiably believe that you have a headache or that it looks to you as if you have hands. To motivate the principle, the explanationist might say something along these lines. If things are going bad for us epistemically, we'll still have unproblematic access to our evidence. (We'd have to have this kind of unproblematic access since having such access is a necessary condition on possessing this evidence.) By appealing to such evidence and relying on inference to best explanation, we can justify beliefs about matters that go beyond this evidence. These matters are things that we could potentially be mistaken about even if our evidence is just the way that it is. ME is designed to capture the intuitive resistance to (S2). There are some matters that we can form justified beliefs about without relying on any further evidence and that is that which the vat operators cannot deceive us about. There are some matters beyond this that we have potentially problematic access to and we can only come to have sufficiently good access to such matters by thinking about the evidence we have to hand.

While the explanationist might avoid the first line of objection (i.e., that it is implausible to think that the justification for believing the Hands-hypothesis must come from inference to best explanation), the maneuvers the explanationist would have to make steer them directly into a second line of objection. The explanationist rejects (S1) and (S2), but accepts the following sufficiency thesis:

> S3: The features of the experiences you have (i.e., (F1)-(F3)) can provide sufficient evidential support for believing that you have hands and that you're not a BIV via inference to best explanation.

If your experiences as of material objects like hands are coherent and regular both within and across sense modalities and there's been a pattern of objects being where you'd expect them to be and appearing as you'd expect them to appear, you can appeal to (F1)-(F3) to justify your belief that you have hands and your belief that you're not a BIV. The view differs from dogmatism insofar as it denies that the one-off experience can do the trick.

A run of experience with certain features is needed to justify a belief, according to the explanationist. The one-off experience isn't going to be fit for purpose. Assuming that your course of experience has the right features, it can provide the right kind of rational support for believing the Hands-hypothesis, provided of course that the experience is of a kind that's better explained by appeal to this hypothesis than a rival hypothesis such as the BIV-hypothesis.

The trouble with the explanationist approach emerges when we think about your relation to (F1)-(F3). These aren't features of any particular, one-off

experience. They couldn't be, the explanationist says. Not only do these features depend upon experiences had at different times, it's important for the explanationist to say that that these features aren't features of any particular, one-off experience since they accept (S3) but reject (S1) and (S2). Since they are features of a *series* of experiences that the subject had at different times, the explanationist has to say that these features aren't epistemically available to you in the way that they'd have to be to justify your beliefs:

The Second Anti-Explanationist Argument

SAE1: The truth of the proposition that your experiences have features (F1)-(F3) depends upon whether certain relations hold between present and past experiences, which means that the truth of the proposition that your experiences have features (F1)-(F3) is incompatible with the BIV-hypothesis.

SAE2: If the truth of the proposition that your experiences have features (F1)-(F3) is incompatible with the BIV-hypothesis, you can only be justified in believing that your experiences have features (F1)-(F3) if such a belief is supported by inference to best explanation.

SAE3: You cannot be justified in believing that your experiences have features (F1)-(F3) by means of an inference to best explanation.

SAEC1: Thus, you cannot be justified in believing that your experiences have features (F1)-(F3).

SAE4: If you cannot be justified in believing that your experiences have (F1)-(F3), the fact that your experiences have these features cannot provide rational support for believing that you have hands or that you're not a BIV.

SAE5: If the fact that your experiences have these features cannot provide rational support for believing that you have hands or that you're not a BIV, you cannot justifiably believe that you have hands or that you're not a BIV.

SAEC2: You cannot justifiably believe that you have hands or that you're not a BIV.

The explanationist has to reject (SAEC2) since that's incompatible with (S3). Unless they can block the argument, they should not say that they are offering us a non-skeptical view.

In support of the first premise, (SAE1), it should be noted that the BIV-hypothesis does not simply target perceptual beliefs about your immediate surroundings, but also beliefs based on apparent memories about the past experiences you take yourself to have. Just as the vat operators have the power to trick a brain in a vat that's stored in a lab into believing that it's sitting on the beach, it can trick a brain in a vat that's been created in a lab only moments ago into thinking that it's been alive for years and that it's had just the sorts of experiences that we do when we see hands during that time. Thus, according to (ME), you cannot justifiably believe your experiences have features (F1)-(F3) unless you can believe this on the basis of an inference to best explanation. Unfortunately, the features of your experiences that you'd need to appeal to in

justifying beliefs about the experiences you've had are largely features of past experiences. They are epistemically off-limits until we have a story about how you could justifiably take your past experiences to have had the relevant features. Thus, so far as your beliefs about what your past experiences were like, it seems that you couldn't justifiably hold such beliefs without relying on inference to best explanation and it seems you don't have the right kind of access to the propositions about the explanandum for that inference to generate any justification.

Once we see why the explanationist doesn't have the right to say that our beliefs about our past experiences are justified, we can see why the explanationist cannot plausibly defend (S3) while rejecting (S1) and (S2). The explanationist needs one-off apparent memories about your past experiences to be the sort of thing that can justify beliefs about these experiences so that you can appeal to features (F1)-(F3) to justify beliefs about the external world, but the explanationist denies that such one-off apparent memories are capable of providing that kind of justification.

Conclusion

The explanationist cannot have it both ways. If the possibility of deception and error are sufficient to show that one-off experiences are incapable of justifying beliefs about external objects, they should be sufficient to show that one-off apparent memories are incapable of justifying beliefs about past experiences. Without such justified beliefs about the character of past experiences, we cannot rationally appeal to features like (F1)-(F3) in the attempt to justify the belief that we're not BIVs. We can't get our hands on the data that we need for inference to best explanation. Thus, the features (F1)-(F3) cannot play any interesting role in any non-skeptical explanationist account of the justification of our beliefs about the external world. If, however, it's possible for one-off apparent memories to justify beliefs about past experiences in spite of the fact that our BIV counterparts could have these apparent memories without having had experiences with features like (F1)-(F3), the explanationist doesn't have any principled reason to think that (S1) and (S2) are mistaken. The skeptical pressures that they think undermine the dogmatist view undermine their view. The explanationist's failure to appreciate this reflects a failure to appreciate the non-trivial problem they face in understanding how we could have the right kind of access to a *series* of experiences for us to appeal to these features in trying to justify our beliefs. If there's never a point at which a particular experience is sufficient on its own to justify a belief that is subject to a skeptical challenge, there is no hope of having any justification for believing things about the external world. [13]

[13] I would like to thank Matthew Frise, Jon Matheson, Kevin McCain, Andrew Moon, and Chris Tucker for discussion and feedback on a previous draft.

References

Beebe, James R. 2009. "The Abductivist Reply to Skepticism." *Philosophy and Phenomenological Research* 79: 605-36.

Black, Tim. 2002. "A Moorean Response to Brain-in-a-Vat Scepticism." *Australasian Journal of Philosophy* 80: 148-63.

BonJour, Laurence. 1999. "Foundationalism and the External World." *Philosophical Perspectives* 13: 229-49.

Feldman, Richard. 2001. "Skeptical Problems, Contextualist Solutions." *Philosophical Studies* 103: 61-85.

____. 2004. "Foundational Beliefs and Empirical Possibilities." *Philosophical Issues* 14: 132-48.

Jackson, Frank. 1977. *Perception: A Representative Theory*. Cambridge: Cambridge University Press.

Kotzen, Matthew. 2012. "Silins' Liberalism." *Philosophical Studies* 159: 61-8.

Lipton, Peter. 2004. *Inference to Best Explanation*. 2nd edition. London & New York: Routledge.

Littlejohn, Clayton. 2012. *Justification and the Truth-Connection*. Cambridge: Cambridge University Press.

Lycan, William. 1988. *Judgment and Justification*. Cambridge: Cambridge University Press.

____. 2002. "Explanation and Epistemology." In *Oxford Handbook of Epistemology*, ed. Paul K. Moser, 408-34. Oxford: Oxford University Press.

McCain, Kevin. 2014. *Evidentialism and Epistemic Justification*. New York: Routledge.

McDowell, John. 1995. "Knowledge and the Internal." *Philosophy and Phenomenological Research* 55: 877–93.

Moser, Paul K. 1989. *Knowledge and Evidence*. Cambridge: Cambridge University Press.

Neta, Ram. 2004. "Skepticism, Abductivism, and the Explanatory Gap." *Philosophical Issues* 14: 296-325.

Pritchard, Duncan. 2012. *Epistemological Disjunctivism*. Oxford: Oxford University Press.

Pryor, James. 2000. "The Skeptic and the Dogmatist." *Nous* 34: 517-49.

Russell, Bertrand. 1912. *The Problems of Philosophy*. Oxford: Oxford University Press.

Silins, Nicholas. 2008. "Basic Justification and the Moorean Response to the Skeptic." In *Oxford Studies in Epistemology 2*, eds. Tamar Szabo Gendler and John Hawthorne, 108-42. Oxford: Oxford University Press.

Vogel, Jonathan. 1990. "Cartesian Skepticism and the Inference to Best Explanation." *Journal of Philosophy* 87: 658-66.

Williamson, Timothy. 2000. *Knowledge and its Limits*. Oxford: Oxford University Press.

Some Problems With Steadfast Strategies for Rational Disagreement

Hamid Vahid

Abstract: Current responses to the question of how one should adjust one's beliefs in response to peer disagreement have, in general, formed a spectrum at one end of which sit the so-called 'conciliatory' views and whose other end is occupied by the 'steadfast' views. While the conciliatory views of disagreement maintain that one is required to make doxastic conciliation when faced with an epistemic peer who holds a different stance on a particular subject, the steadfast views allow us to maintain our confidence in our relevant beliefs. My aim in this paper is not to adjudicate between these views. Rather, I shall focus on a particular strategy, namely, denying the appearance of epistemic symmetry between peers, that the steadfast views standardly invoke in support of their position. Having closely examined certain representative examples of the steadfast approach, I will argue that this strategy is fundamentally flawed.

Keywords: rational disagreement, 'conciliatory' views, 'steadfast' views, epistemic norms, epistemic peer

It is common for equally competent epistemic agents to arrive at different conclusions in regard to a particular subject matter on the basis of roughly the same body of evidence. Although it is rarely possible for two people to share all their evidence, let us, idealizing away from the actual cases, call two agents, S_1 and S_2, epistemic peers if they are equal with respect to the general intellectual virtues and are acquainted with the same evidence E in regard to a particular proposition p. Suppose further that, having reached different conclusions on the basis of E, they are then apprised of each other's competing views about p but continue to disagree. The question of what epistemic peers ought to do under such circumstances constitutes the main concern of the epistemology of disagreement.

There are roughly two fundamental positions on the proper epistemic response to a situation involving peer disagreement. According to the so-called 'conciliatory' views, S_1 and S_2 should revise their attitudes towards p, though the extent of such revisions varies depending on how conciliatory the views in question are.[1] In the opposite direction, one finds the so-called 'steadfast' views which require peers to maintain their confidence in their attitudes in the face of

[1] See, for example, Feldman 2006, Christensen 2007 and Elga 2007.

© Symposion, 1, 1 (2014): 89-107

disagreement.[2] My aim in this paper is not to adjudicate between these views. Rather, I wish to focus on a strategy that the steadfast views standardly invoke in support of their position and against conciliationism. Having closely examined certain representative examples of the steadfast approach, I will try to show that this strategy is fundamentally flawed.

1. Reasonable Disagreement after Full Disclosure: The Questions

Consider again our agents S_1 and S_2 who hold different attitudes towards a proposition p on the basis of the evidence E. Suppose while S_1 believes that p is true, S_2 believes that it is false. The fundamental question is whether they can reasonably maintain their beliefs after becoming fully aware of each other's views. Following Feldman (2006; 2007), it is important to realize that the question of the possibility of reasonable disagreement actually boils down to two questions.

(a) Can both S_1 and S_2 reasonably maintain their beliefs after learning of the other's opinion?

(b) Can both S_1 and S_2 reasonably continue to disagree after disclosure while at the same time think that the other's belief is reasonable as well?

To say yes to the first question is to admit that reasonable disagreement after full disclosure is possible, while a positive answer to the second question is tantamount to admitting that it is also possible to have *mutually* recognized disagreement. It is arguable that intellectual humility requires that the steadfast views have the resources to answer the second question in the positive. It is, thus, the question (b) that is going to feature prominently in our discussion of the steadfast views.

1.1. Steadfast Views: The Standard Strategy

As noted earlier, there is a major divide separating positions on the issue of the possibility of reasonable disagreement. There are, on the one hand, the conciliatory views that require epistemic peers to temper their views in the face of disagreement and there are, on the other hand, the steadfast views that allow peers to maintain confidence in their beliefs after disclosure. The steadfast views are, however, initially puzzling for it is difficult to see how peers who are recognized to be roughly equal in terms of intellectual virtues can reach and maintain incompatible, and yet equally reasonable, beliefs on the basis of the same body of evidence. If the peer relationship involves evidential as well as general intellectual equality, how can S_1 and S_2 come to adopt incompatible attitudes towards a particular proposition p? Moreover, how can they reasonably maintain their respective attitudes when, after full disclosure, they come to be apprised of all their evidence, background beliefs, etc.? What is it that

[2] See, for example, Kelly 2005, Bergmann 2009 and Goldman 2010.

gives, S_1, for example, the right to maintain her belief and resist the epistemic pull of S_2's contrary opinion? The same can be asked of S_2 and her belief.

A natural strategy (henceforth, 'the standard strategy'), for the steadfast view, is to deny the appearance of epistemic symmetry between peers. However, given our assumptions, any putative asymmetry must be characterized in terms other than the agents' general intellectual virtues or their normal body of evidence. It might, for example, be claimed[3] that S_1 possesses insights about the disputed proposition that S_2 lacks or that S_1 and S_2 are employing different systems of epistemic norms (Goldman 2010).[4] Whatever the merits of these suggestions, what is important to note is that no steadfast account of the possibility of reasonable disagreement can come for free. The initial epistemic symmetry between peers has to be broken at some point along the way. If the parties to a dispute are to maintain different attitudes in regard to a particular topic this should be traced back to some prior difference on their part. It is difficult to see how, despite the full epistemic symmetry, peers can rationally arrive at different conclusions in regard to a particular subject matter. To see this, let us consider a recent steadfast account of the possibility of reasonable disagreement due to Marc Moffett (2007).

While agreeing that the most plausible line of defense for a steadfast response to our central question is to appeal to some sort of epistemic difference on the part of the peers, Moffett claims that it is possible to avoid downgrading their competence while maintaining that they can reasonably hold on to their views. The trick is "to take advantage of the transient underdetermination of theory by evidence" (Moffet, 2007, 360). It is not, however, clear why the postulation of an epistemic asymmetry between peers is tantamount to downgrading their competence. To assume that S_1 and S_2 are equal with respect to general intellectual virtues is consistent with the possibility of one of them being more competent or reliable on a particular topic. More importantly, however, the appeal to the underdetermination thesis hardly explains the claim that the same body of evidence (E) can rationalize two contradictory claims p and not-p. If E supports p (to a sufficient degree), then it would be reasonable to believe p on that basis. But this would, in turn, entail that it is not reasonable to believe not-p. It is therefore not clear how the same body of evidence can rationalize two contradictory beliefs. To say that the beliefs (theories) are underdetermined by E is merely to repeat the claim in question. The underdetermination thesis is, in other words, semantically too close to the claim that reasonable disagreement is possible to explain it.

Moreover, assuming that the underdetermination thesis is true, it can, at most, explain how reasonable disagreement is possible before disclosure (i.e., in isolation). It does not tell us why S_1 and S_2 can maintain confidence in their

[3] For example, van Inwagen 1996 and Bergmann 2009.

[4] For a different account of asymmetry in terms of the information one has about onself and one's opponent see Sosa 2010.

beliefs *after* disclosure. After all, the conciliationists claim that when S_1 and S_2 learn that the other holds a different attitude, on the basis of the same evidence, the awareness of this fact would provide a defeater for their beliefs.

In response, Moffett appeals to Sklar's principle of methodological conservatism (MC) according to which it would be reasonable to hold on to a hypothesis despite coming to know of evidentially equivalent alternatives (Sklar 1975). But (MC) is precisely what the concilaitionists are challenging. Unless Moffett can provide some independent justification for (MC), he cannot, on pain of begging the question against the conciliationists, appeal to it to defend his steadfast view. He does, however, make some remarks in defense of (MC) by referring to Harman's (1986) observation that some such conservative principle is indispensable to our epistemic well-being for

> if we are not able to rely on our standing beliefs without re-verifying them or – what is cognitively impossible for creatures like us – carrying along our justifications, we will be unable to make justifiable progress in our theoretical undertakings (Moffet 2007, 361).

But Harman's observation actually concerns a different version of the thesis of epistemic conservatism according to which one is justified to continue to hold a belief as long as there are no good reasons against it. What underpins Harman's principle of conservatism is the phenomenon that people do not usually keep track of the justification relations among their beliefs which is presumably why they tend to preserve their beliefs in the face of evidential discrediting.[5] These considerations hardly lend any support to (MC).

Moffett seems to realize that Harman's principle is not sufficient to establish (MC) and, thus, his version of the steadfast view. He, thus, chooses to highlight, what he takes to be, the implausible skeptical consequences of the conciliatory views (like Feldman's), namely, that they would be 'epistemically devastating' if consistently applied across the board. But this remark is hardly relevant to our concerns here, not least because not all versions of conciliationism have such implications. Moffett's animadversions, thus, seem to lead nowhere. However, he makes one final observation in defense of his view claiming that holding to one's beliefs has certain cognitive advantages in that it helps us develop and refine theoretical frameworks and worldviews.

But are such alleged advantages to be counted as *epistemic* reasons for beliefs? Yes, he thinks, unless one

> [construes] the epistemic in an inappropriately narrow manner, as reasons which the belief itself is more likely to be true (Moffett 2007, 363).

But this revisionist step threatens to change the rules of the game, for it is no longer clear that he can continue to characterize an individual x as an

[5] For an analysis of Harman's version of epistemic conservatism see Vahid 2004.

> epistemic peer with respect to a given domain (if and) only if we regard x's judgments concerning that domain to be as likely as our own to be correct given the same evidence (Moffett 2007, 356).

More importantly, for our purposes here, the appeal to the pragmatic virtues of belief perseverance as reasons for why epistemic peers can hold on to their beliefs in the face of disagreement is to postulate an asymmetry, though not in the space of reasons but, in the space of pragmatic considerations. Just as the appeal to insights might help justify why epistemic peers can reasonably hold on to their beliefs after disclosure, so does the appeal to the purported pragmatic consequences of maintaining such beliefs. Either way, no steadfast view comes for free as, one way or another, it is forced to appeal to factors regarding which epistemic peers diverge. Let us then return to our main question and examine some steadfast views that employ the standard strategy to show how reasonable disagreement is possible. I begin with a recent account (due to Goldman 2010) that locates the required asymmetry in the systems of epistemic norms that are said to license the agents' beliefs.

2. Objectivity-based Relativism as the Basis of Reasonable Disagreement

To show that reasonable disagreement is possible, Goldman appeals to two fundamental theses of epistemic objectivism and epistemic relativism. He takes 'objectivism' to mean that there is a uniquely right epistemic system (E-system) comprising of rules and norms directed at doxastic attitudes. Thus, according to objectivism, for any proposition p, a right E-system determines what attitude someone should adopt towards p on the basis of some evidence e. The thesis of epistemic relativism, on the other hand, construes justification statements as covertly referring to the E-system the speaker accepts. The combination of these elements, Goldman thinks, would provide a solid basis for the possibility of reasonable disagreement. To see this, consider our two agents S_1 and S_2 believing p and not-p respectively on the basis of the same evidence E. Suppose, in accordance with Goldman's objectivity-based-relativism (OBR-) account of justification, S_1 and S_2 operate with two different epistemic systems, E_1-system and E_2-system respectively that authorize different attitudes towards p. Since, by hypothesis, there is a uniquely correct E-system, this means that S_1 and S_2 cannot be both objectively (O-)justified in their beliefs. So, at the level of the target proposition p, at least one of those agents is unreasonable or unjustified. Does this mean that the story for the possibility of reasonable disagreement is over? No, for Goldman thinks that his OBR-account of epistemic justification has the resources to show how S_1 and S_2 can reasonably maintain their beliefs on the basis of the same evidence E.

To show this, Goldman starts by claiming that S_1 and S_2 are both O-justified in believing their respective E-systems or individual E-norms. To defend this claim, he asks us to consider communities and cultures where young children and people are instructed by their elders to believe and follow certain

E-norms like “If the Scripture says p, you should believe that p” or “If scientists agree on p, you should assign a high credence to p.” He further assumes that these "children hear roughly the same testimony from numerous elders and no conflicting testimony" (Goldman 2010, 198). Such scenarios, he thinks, provide examples in which an incorrect E-norm is O-justifiably believed to be correct. These norms would, in turn, sanction incompatible attitudes in regard to particular propositions. How does this bear on the question of whether reasonable disagreement is possible? According to Goldman, since S_1 and S_2 are O-justified in believing the E_1-norm and E_2-norm respectively, they may also enjoy a distinct but significant justificational status for the target beliefs those norms authorize. So while they may not be first-order O-justified in their beliefs, they are, nonetheless, iteratively O-justified in holding those beliefs, that is, we can say of, say, S_1 that “S_1 is O-justified in believing that she is O-justified in believing p.”

Likewise for S_2. Although second-order O-justifiedness does not entail first-order O-justifiedness, it does, says Goldman, make a significant contribution towards the reasonability of a first-order belief as he denies that the first-order justificational status of an attitude fixes its overall reasonability. By bringing epistemic norms to bear on the reasonability of the target beliefs, Goldman admits that he is, in effect, claiming that S_1's and S_2's total evidence differ. For although they have, what he calls, the same ‘material’ evidence (E), their norm evidence (E_1/E_2-norms) differ which is why their attitudes towards p can legitimately diverge. Thus, an OBR-account of epistemic justification, he concludes, can explain how divergent attitudes based on the same body of evidence can be both reasonable.

There are, however, a number of places in Goldman's argument for the possibility of reasonable of disagreement where one can take issue with it. By appealing to an OBR-account of justification, Goldman is obviously implementing the standard strategy to defend his version of the steadfast view. He only chooses to construe the required epistemic asymmetry in terms of different epistemic norms that guide the agents' doxastic behavior. Thus, while S_1 adopts the E_1-norm to arrive at p, S_2 adopts the E_2-norm to arrive at not-p. To assess Goldman's account of the possibility of reasonable disagreement, I propose to examine it in view of the following questions: (1) Is Goldman's account internally coherent? And (2) does it successfully explain how reasonable disagreement is possible? I begin with the first question.

What makes Goldman say that the epistemic agents (like the children of his example) are O-justified in believing incompatible E-systems or E-norms? Presumably, this is because they adopt their E-norms on the basis of different testimonial evidence coming from their elders in different communities. Consonant with his norm-based account of epistemic justification, Goldman thinks that S_1's and S_2's O-justified beliefs in their pertinent E-systems themselves also involve norms (though of a more fundamental kind). These

fundamental norms are source authorizations like the following 'generic testimony-based norm' pertinent to his example involving children: "If a random speaker or writer testifies that p, then (in the absence of defeating conditions) the agent should believe that p." So there are actually two kinds of evidence involved here that ground S_1's and S_2's beliefs in their incompatible E-norms. There is the norm evidence, like the generic testimony-based norm above, and there is the material evidence consisting of the different testimonies that the children obtain in the form of instructions form their elders.

Suppose we agree that, on the basis of their total evidence, S_1 and S_2 are justified in believing that their respective E-norms are correct. On the basis of these norms, S_1 and S_2 then go on to hold different attitudes towards p even though they happen to share the same material evidence E. As we have seen, however, Goldman denies that S_1 and S_2 can be both O-justified in holding different attitudes towards the target proposition p. But it is not clear why S_1 and S_2 can be O-justified at the norm level but not O-justified at the target-belief level. That is, it is not clear why S_1 and S_2 can be O-justified in holding different attitudes about the correct E-system but fail to be O-justified in holding incompatible attitudes towards p. For, on both levels, they have the same type of total evidence consisting of appropriate norm evidence as well material evidence. The only difference seems that, at the norm level, it is the norm evidence that is shared by S_1 and S_2 with their material evidence varying while, at the target-proposition level, it is the material evidence E that is fixed with the norm evidence being different (E_1/E_2-norms). So if S_1 and S_2 can be both O-justified at the norm level because their total evidence differs, there is no reason why they cannot also be O-justified at the target-proposition level. Indeed, if, as Goldman claims, S_1 and S_2 are both O-justified in believing the correctness of the E_1-norm and the E_2-norm respectively, then, obviously enough, they should also be O-justified in believing what these norms authorize. A belief that is grounded in another belief that is O-justified is itself O-justified.

It might be argued that what grounds the distinctness of the epistemic status of S_1 and S_2 at these two levels has to do with the fact that the E_1-norm and the E_2-norm cannot both belong to the uniquely correct E-system. At least one of these norms must be incorrect and so must the belief it sanctions. But this point is irrelevant to what is at issue here which is justification not truth. A belief can be false and yet O-justified. What counts for justification at both the norm and the target-proposition levels is one's total evidence which might include misleading evidence. As Goldman himself emphasizes, truth is not necessary for justification:

> what determines a belief's reasonability is the agent's evidence…The same point holds on the topic of norm correctness. The actual rightness of an E-system does not determine the reasonability of an agent's conforming to it. What is critical is the agent's evidence about its rightness. If an agent conforms her attitude to the prescriptions of a properly chosen E-system, this should be an important – perhaps decisive – element in assessing the attitude's

> reasonability, even if the evidence supporting that E-system's rightness happens to be misleading (Goldman 2010, 206).

Accordingly, since both S_1 and S_2 have evidence for their adopted E-norms (which is why they are said to be both O-justified in believing them), they are also O-justified in holding the attitudes that those norms license about p. So, they can also be O-justified at the level of the target proposition despite having misleading evidence.

I think Goldman's inconsistent treatment of the agents' beliefs at the norm and the target-proposition levels arises from his expressed wish to combine elements form both objectivism and relativism in his theory of epistemic justification. On the one hand, the objectivist ingredient prompts him to link

> justifiedness not to any random E-system but to a *right* E-system, because…[o]nly a right epistemic system has the appropriate connection with objective justifiedness (Goldman 2010, 193)

or reasonability. This, in turn, leads him to deny that S_1 and S_2 can both be O-justified in holding different attitudes at the target-proposition level. The relativist ingredient, on the other hand, impels him to say that the

> actual rightness of an E-system does not determine the reasonability of an agent's conforming to it (Goldman 2010, 206)

which is presumably why he claims that S_1 and S_2 are both O-justified in adopting different E-systems. Let us now turn to our main question of whether Goldman's OBR-account of epistemic justification can successfully explain the possibility of reasonable disagreement. We need to address this question both before and after the agents come to know of their respective views. I start with the possibility of reasonable disagreement before full disclosure.

According to Goldman, reasonable disagreement before disclosure is possible because S_1 and S_2 are iteratively, though not first-order, justified in holding different attitudes towards p. This contention is, as we have seen, in turn, grounded in Goldman's claim that the agents are O-justified in believing that their incompatible E-norms are correct. There are, however, a number of ways that this latter claim can be challenged. Recall Goldman's example of the children in isolated communities who adopt incompatible E-norms on the basis of their elders' advice. Now, given Goldman's 'reliabilist criterion of [E-]system superiority' (Goldman 2010, 194), shouldn't those testimonial sources themselves be reliable if the children are to be *objectively* (O-)justified in believing their E-norms? One may concede that the children enjoy some sort of justification for their beliefs but objective justification seems to require that their evidence be, at least, reliably adequate.

Indeed, Goldman's examples resemble the so-called 'epistemic poverty cases' in the philosophical literature which are often invoked to show the inadequacy of deontological justification. To say that a belief p is deontologically justified is to say that in holding that belief, an agent has flouted no (subjective)

epistemic obligations, and is, thus, subject to no blame or disapproval. A standard way to express the thought that an agent is deontologically justified in believing p is to say that he is iteratively justified (in a truth-conducive sense) in holding that belief.[6] The deontological conception of justification has been criticized for, among other things, failing to give us what we expect of epistemic justification. The thought behind this criticism is often formulated in terms of the so-called 'epistemic poverty cases' where, despite doing all that can be reasonably expected of agents, they form their beliefs on less than adequate grounds. One such case, 'the cultural isolation case' (Alston 1988), involves an agent, S, growing up in an isolated community in which everyone unquestionably accepts the traditions of the tribe as authoritative. Having had no opportunity to come across circumstances in which this authority is challenged, S can hardly be blamed for holding beliefs that are grounded in the traditions. He is thus deontologically justified in holding those beliefs despite the beliefs being unjustified (in a truth-conducive sense).

The epistemic circumstances of S are similar to those of the children in Goldman's scenario who are said to be exposed to "no conflicting testimon[ies]." It, thus, seems unreasonable to describe the children in Goldman's example as having objectively (O-)justified belief in their epistemic norms. Indeed, Goldman himself, when challenged on this issue, goes on to concede that he can be taken to be addressing a conception of "justification [that is] grounded entirely in one's subjective perspective" (Goldman 2010, 199, fn.9). Of course, once we grant that the agents are only iteratively justified in believing their adopted E-norms, it would be quite plausible to think (with Goldman) that they are also iteratively justified in holding beliefs that those E-norms sanction. In this case there would be no inconsistency in the treatment of beliefs at the norm and the target-proposition levels. However, as we can see, this requires giving up on the notion of objectivity that Goldman's account involves.

Another reason for denying objective justifiedness at the norm level is this: Any explanation of the possibility of reasonable disagreement between epistemic peers that helps itself with the assumption that they are O-justified in believing their respective E-norms would be question-begging. For suppose we assume, with Goldman, that S_1 and S_2 are O-justified in believing incompatible E-norms. To say this, however, is to admit that reasonable disagreement is possible. So, on pain of circularity, this assumption cannot be used as a premise (as in Goldman's account) to justify the conclusion that reasonable disagreement is possible. Differently put, Goldman manages to show that reasonable disagreement is possible at the level of the target proposition only by assuming it at another level, namely, the norm level.

Finally, suppose we agree, with Goldman, that S_1 and S_2 are only iteratively justified in holding different beliefs about p. Does this settle the issue of

[6] See, for example, Alston 1988. Goldman himself uses the term 'weak justification' to describe this conception of justification. See, for example Goldman 1988.

reasonable disagreement before closure? No, because the notion of iterative justifiedness is too weak. Indeed, one could already acknowledge this fact about the agents without buying into Goldman's conceptual framework. To say, for example, that S_1 is iteratively justified in believing p is, as we have seen, to say that she is deontologically justified in holding that belief and that she has flouted none of her epistemic obligations in forming that belief. In other words, it is to say that her belief is blameless and responsibly formed. But this is something one could already acknowledge about such agents having assumed that they are rational peers.

Let us now see how Goldman's observations bear on the question of the possibility of disagreement *after* full disclosure. Hardly at all, it seems. Indeed, by the end of his discussion, Goldman confesses that his approach is only intended to address the issue of reasonable disagreement in isolation (or 'synchronic disagreement' as he calls it). It is not concerned with the 'diachronic' question, namely, with how an agent should epistemically react when, having formed her belief at t, she comes to know of her peer's contrary opinion at a later time t'. I end my discussion of Goldman, however, by arguing that not only his account is silent about whether one can reasonably maintain her belief after disclosure, it *cannot* recognize such a possibility.

To see this, we may recall that, to argue for the possibility of reasonable disagreement at the level of the target proposition, Goldman needed the premise that it is possible for agents to be O-justified at the norm level. The argument for *this* premise, on the other hand, consisted of Goldman's presenting a case, "a likely scenario in many communities both historical and contemporary" (Goldman 2010, 198) where, he claims, it is easy to see how agents can be O-justified in believing incompatible E-norms. The case involved children living in communities where they follow their elders and teachers about certain matters without much critical discussion and without being exposed to "conflicting testimon[ies]" (Goldman 2010, 198). What this implies is that intellectual or epistemic isolation is necessary if Goldman's scenario is to give him the premise he wants. However, with such a requirement in force, Goldman's agents can hardly proceed to a state of full disclosure. For, after disclosure, when their intellectual isolation ends, they will no longer be entitled to having O-justified beliefs in their E-norms, and, thus, there would be nothing that would authorize them to have different attitudes towards the target proposition p. After full disclosure, there will be no opportunity to maintain reasonable disagreement. In view of its structure, Goldman's account of reasonable disagreement is bound to crumble in the state of full disclosure. I shall now turn to another defense of the steadfast view (due to Bergmann 2009) that also appeals to the standard strategy.

3. Insights and Error Theories as the Bases of Reasonable Disagreement

Right from the start, Bergmann states that his target is to explain how reasonable disagreement is possible after full disclosure. Following Plantinga, he distinguishes between two kinds of rationality, internal and external rationality. A belief is internally rational if and only if it is an epistemically appropriate response to the agent's evidence (consisting of her mental states). Internal rationality, thus, concerns the epistemic standing of a belief downstream from experience. One might, thus, say of the beliefs of a brain a in a vat (BIV) or a victim of the Cartesian demon that they are internally rational or justified. External rationality, on the other hand, requires that, in addition, the agent's cognitive mechanism be functioning properly. This means that since the BIV's experiential evidence is due to malfunction, her beliefs are not externally rational.

Bergmann then appeals to the standard strategy to show that, whichever conception of rationality is in place, mutually recognized reasonable disagreement is possible. Unlike the previous steadfast views, he does not construe the required epistemic asymmetry in terms of pragmatic considerations or epistemic norms. Rather, he unpacks it in terms of the insights that agents might possess. Thus, we might say of two agents, S_1 and S_2, that they differ not only with respect to their attitudes towards p but also in regard to, what Bergmann calls, their broader outlooks O_1 and O_2 which he takes to include the following ingredients.

O_1 • p

• a theory of error (applied to epistemic peers who believe the key ingredients of O_2) according to which the apparent insight that underpins the key ingredients of O_2 is not a genuine insight

O_2 • not-p

• a theory of error (applied to epistemic peers who believe the key ingredients of O_1) according to which the apparent insight that underpins the key ingredients of O_1 is not a genuine insight

So before S_1 learns of S_2's contrary view, she has apparent insights that both p and the error theory in O_1 are true. The right response to these insights is to believe p and the associated error theory. Thus, both these beliefs are, by definition, internally rational. We can further stipulate, says Bergmann, that, given the strength of these apparent insights, the beliefs in question are very strongly justified. This takes care of the question whether the parties to the dispute can rationally disagree before disclosure. But what happens after S_1 learns of her disagreement with S_2 (whom S_1 takes to be an epistemic peer)? Let us say that the only piece of evidence that is added to S_1's body of evidence after disclosure is S_2's report of her apparent insight in support of her belief that not-p. Does S_1's recognition of this fact count as a defeater for S_1's own belief that p?

No, says Bergmann, because, given S_1's rational high confidence in her error theory for those peers (like S_2) who reject p, S_1 can rationally hold on to her belief: "As a result S_1 remains internally rational in her belief that p despite her recognition that S_2…believes not-p" (Bergmann 2009, 340). The same thing can, of course, be said of S_2's belief that not-p. Bergmann further claims that these remarks also explain why S_1 and S_2 can both recognize that the other is internally rational in holding her respective belief. So mutually recognized reasonable disagreement is also possible.

Let us now examine to what extent these observations succeed in explaining the possibility of reasonable disagreement after disclosure. I shall focus on the notion of internal rationality that Bergmann takes to be the more fundamental notion and, in fact, equivalent to justification. So, is reasonable disagreement possible when it is internal rationality that constitutes the epistemic status of beliefs? Again, the full answer to this question requires not only the investigation of the possibility of reasonable disagreement after full disclosure but also the possibility of the mutually recognized variety.

Recall that a belief p is said to be internally rational if and only if it is an appropriate epistemic response to the agent's evidence (mental states) regardless of whether or not this experiential evidence is the result of cognitive malfunction on the part of the agent's cognitive mechanisms. Let us concede that S_1's and S_2's beliefs are internally rational before disclosure. The question is whether they also remain internally rational after full disclosure. After disclosure, S_1's and S_2's bodies of evidence expand to take into account what each learns about the other's view about p. S_1 then has, in her body of evidence, not only her insight for the truth of p as well as the associated error theory but also the belief that S_2 has an apparent insight that not-p. According to Bergmann, given the strength of S_1's apparent insight for the belief p and for her error theory, her recognition of S_2's insight for not-p will fail to undermine her convictions and so she will remain internally rational in believing that p. In other words, S_1's belief that p remains an epistemically appropriate response to her mental states after full disclosure. Likewise for S_2's belief that not-p. So reasonable disagreement after full disclosure is possible.

There is, however, a slight complication here. To argue for this conclusion, Bergmann needs to assume that S_1 and S_2 maintain 'rational high confidence in their respective error theories' after full disclosure. For it is only because they believe their respective error theories with such high confidence that each is able to resist the negative epistemic influence of her newly found evidence that the other holds a strong insight for an incompatible belief and, thus, remain internally rational in her attitude towards p:

> [T]hese beliefs – of S_1 in…the error theory of O_1 and of S_2 in…the error theory in O_2 are each partially based on apparent insight that the propositional content of the belief so based is true (Bergmann 2009, 339).

So S_1's and S_2's beliefs in their respective error theories are internally rational. They are, on the other hand, incompatible as they pass different judgments on what counts as a genuine insight about p. Taken together, what these remarks suggest is that S_1 and S_2 can reasonably disagree over what content a genuine insight about p must have. This means that, as with Goldman's account, it is only by assuming that reasonable disagreement is possible at one level (the level of error theories) that Bergmann is able to show that it possible at the level of the target proposition p.

Of course, Bergmann does not deny that sometimes recognized disagreement can provide one with a defeater of one's belief. Under such circumstances reasonable disagreement is not possible as peers are forced to temper their views. But what are these circumstances? According to Bergmann, these are circumstances in which one *should* disbelieve or seriously doubt that (C) "one is, on this occasion, more trustworthy than one's peer who holds a different attitude towards p." However, he refrains from giving a general recipe as to when an attitude of disbelief or doubt is the epistemically appropriate response to one's mental states (after full disclosure). Each case, he says, should be examined individually.

But it is not clear why the recognition of peer disagreement should be thought to undermine, say, S_1's or S_2's confidence in (C), that is, in believing that each is more trustworthy than the other with respect to p. (C) is what the error theories in their outlooks entail. When, for example, S_1 believes on the basis of her error theory that the insight supporting S_2's belief is not genuine, she would consequently take herself to be more trustworthy than S_2 on the topic of p. Given Bergmann's own arguments, S_1's recognition of S_2's disagreement about p should not undermine her belief in her comparative reliability. If S_1 *is* to be vulnerable to S_2's adverse influence, one should make sure that she is no longer protected by her error theory. Once she loses her epistemic immunity, her beliefs will become vulnerable to the adverse influence of peer disagreement. But to deny that she is entitled to her error theory in such circumstances is to hold that peers cannot reasonably hold on to their beliefs in their incompatible error theories under such conditions. Once again, one has to assume, though in the opposite direction, that peer disagreement is *not* possible at one level (the level of error theories) to show that it is *not* possible at another level (the level of the target proposition).[7]

[7] Interestingly, this seems to be what happens in Feldman's 2006 conciliationist argument against the possibility of reasonable disagreement. Feldman denies that one can appeal to the different insights that the agents might have to explain the possibility of reasonable disagreement after full disclosure. If, prior to disclosure, S_1 had good reason to think her belief that p is well supported, after disclosure, says Feldman, S_1's expanded body of evidence would no longer support that belief. Rather, "it makes suspending judgment on this matter [' S_1's belief that p is well supported'] the reasonable attitude" (232). It follows, Feldman claims, that the expanded body of evidence would no longer support p either because "if still does support p, then it supports [S_1] reasonably having a complex attitude that she would express as

Before I turn to, what I take to be, a fundamental problem with Bergmann's defense of the steadfast view, it would be worthwhile to note another dimension with respect to which his account resembles Goldman's. We noted that, according to Goldman, S_1 and S_2 can reasonably disagree with each other because, though not first-order justified, they are nevertheless iteratively justified in holding different attitudes towards the target proposition p. This does not seem to be very different from what Bergmann claims about the internal rationality of those agents after disclosure. Although a great deal needs to be said about the notion of internal rationality and how it differs from external rationality, Bergmann's sketchy remarks and his examples can still give us a rough idea of how the distinction is to be understood. An internally rational belief, he says, has to do with what goes on in belief formation 'downstream from experience' regardless of whether or not the experience itself is due to cognitive malfunction. He cites, as examples of internally rational beliefs, the beliefs of the victims of an evil demon or a superscientist. As we have seen, however, such beliefs are often construed as involving iteratively justified beliefs (involving a truth-conducive conception of epistemic justification). Goldman himself calls such beliefs 'weakly,' as opposed to 'strongly,' justified.[8] Weak justification does not imply strong justification just as, according to Bergmann, internal rationality is distinct from external rationality. So, at most, what follows from Bergmann's account is that S_1's and S_2's beliefs are, after disclosure, responsibly formed and non-culpable. But, as with Goldman's account, this is a fact that one could already acknowledge about S_1 and S_2, having recognized them as rational and responsible peers.

Finally, there is something about Bergmann's way of implementing the standard strategy that threatens to violate, what I take to be an adequacy condition on an acceptable theory of reasonable disagreement. The idea is that an epistemology of disagreement must be sensitive to the distinction between reasonable disagreement in isolation and reasonable disagreement after full disclosure. Suppose S_1 comes to believe that p as it seems to her that her evidence supports p. Let us further assume that since S_1 has carefully examined her body of evidence and there is no countervailing evidence against p, S_1 is justified in believing that p. Being aware of her fallibility and her own history belief formation, S_1 can obviously conceive of others contradicting her belief on the basis of the same body of evidence. The mere conceivability of disagreement, however, is no bar to S_1's maintaining her belief if the belief has been responsibly formed. We may thus grant the rationality of S_1's belief in isolation. Suppose now S_1 learns that S_2 has come to believe not-p on the basis of the same body of

follows: I believe p, but I suspend judgment on whether my evidence supports p" (232). Whatever the merits of Feldman's argument, it shows the impossibility of reasonable disagreement at the level of the target proposition only by assuming it another level, namely, for the belief that the belief that p is well supported.

[8] See footnote 6.

evidence. S_1 realizes that at least one of them is in error in their evaluation of the evidence. Who is to blame? Given that the interesting cases of disagreement are those where it is not obvious which belief the evidence really supports, S_1 and S_2 are in a symmetrical position in so far as the question of blame is concerned.[9] It would not thus be acceptable if S_1 were to think that the error lies with S_2 (simply because it is her own belief that is at stake) without *explaining* why this is so.

To condone S_1's attitude is to fail to attach any epistemic significance to the distinction between reasonable disagreement in isolation and reasonable disagreement after full disclosure. The central question in the epistemology of disagreement is what one should do with one's own belief when one learns of an epistemic peer's contrary opinion. So I take it that any steadfast view of reasonable disagreement that claims that one can reasonably maintain ones' belief in the face of disagreement merely on the a priori ground that views incompatible with one's own are flawed and ought to be ignored is violating an adequacy condition on an acceptable theory of reasonable disagreement. In other words, any account of why it is reasonable to stick to one's view after full disclosure which rules out the information about peer disagreement as relevant merely on the ground that views contrary to one's own are in error is inadequate. This conclusion can be reinforced by noting that this way of securing reasonable disagreement would also undermine the epistemic significance between possible and actual disagreement. For what is of importance for the question of the possibility of reasonable disagreement is actual, rather than, possible peer disagreement. As noted above, the fact that one's belief might possibly be contradicted by similarly intelligent and intellectually virtuous people merely reflects our fallible and imperfect epistemic predicament. Now, it seems to me that that, by incorporating error theories in his theory of epistemic justification (rationality), Bergmann's defense of the steadfast view violates the aforementioned adequacy condition on an acceptable theory of reasonable disagreement.

To clarify further, suppose, having argued for the view that free will is incompatible with determinism, van Inwagen learns of David Lewis's contrary opinion. *After* carefully examining Lewis's reasons, he decides that Lewis is in error. To explain, he claims that he has insights that are denied to Lewis. Now, while this might be, methodologically speaking, a legitimate thing to do if one wishes to defend a steadfast position regarding a particular dispute, it raises problems of the sorts mentioned above when it is turned into a general strategy for defending the steadfast views. For to show that agents can reasonably maintain their views in the face of disagreement, we must make sure that their cognitive systems are armed, as it were, with something like defense mechanisms that are, ceteris paribus, automatically activated whenever the agents' beliefs become vulnerable to adverse epistemic influence as in the peer

[9] See Feldman 2006.

disagreement cases. Such built-in defense mechanisms tend to epistemically immunize the agents against possible threats to their cognitive systems. This way an agent will be allowed to maintain her belief whenever she is in an epistemically hostile situation. Indeed, given such built-in defense mechanisms, the actual cases of disagreement would be as innocuous for one's views as are the possible ones.

Now, this seems to be what Bergmann is doing with his postulation of error theories as essential ingredients of the agents' outlooks. These error theories are formulated in terms of the contrary views of the agents' peers. For example, S_1's error theory rejects as not genuine the insight of any possible peer who believes not-p and who subscribes to another incompatible error theory. Thus, when S_1 learns that there *is* an epistemic peer, S_2, who actually believes the ingredients of O_2, her cognitive defense mechanism automatically starts to kick in to protect her against the adverse epistemic influence of S_2. Although, S_1's error theory does not particularly target S_2, it is intended to apply, as we have seen, to "those roughly equal in intellectual virtue who…reject p" (Bergmann 2009, 339). This is why, the news of the S_2's disagreement would hardly catch S_1 by surprise since given "S_1's rational high confidence in her theory of error," her recognition of S_2's belief that not-p would not count as a "fairly powerful potential defeater" for S_1's belief that p (Bergmann 2009, 340).[10] S_1's error theory would entitle S_1 to her belief at any time regardless of whether or not she has learned of any peer disagreement. Accordingly, on Bergmann's steadfast view, reasonable disagreement in isolation and reasonable disagreement after disclosure are rendered epistemically on a par. In so far as the agents are able to maintain their confidence in their error theories they need not worry about whether or not others disagree with them.

Another factor that ensures the irrelevance of the distinction between reasonable disagreement in isolation and reasonable disagreement after disclosure, on Bergmann's account, concerns the type of evidence that he takes to be germane to the rationality of one's beliefs. He takes insights to be what support the key ingredients of the agents' outlooks. Insights are, however, not the sort of things that can be fully shared:

> [E]ven after full disclosure, S_1 and S_2 do not have the same evidence. Reporting their apparent insights to each other is not the same as giving those apparent insights to each other (Bergmann 2009, 339).

So, given that all the new evidence that S_1 and S_2 might obtain after full disclosure are the *reports* of their insights, the strength of their evidential situations could hardly change after they learn of each other's views. This would

[10] I think the same problem afflicts Goldman's OBR-account of reasonable disagreement. Once epistemic justification is taken to be norm-relative, agents who subscribe to a particular set of E-norms can safely ignore the views of those peers who adopt different E-norms and regard them as unjustified.

provide another reason why, on Bergmann's account, it hardly matters for determining the reasonability an agent's' view whether she is considered in isolation or after exposure to peer disagreement. I conclude therefore that Bergman's account fails to respect an adequacy condition on acceptable theories of reasonable disagreement.[11]

Conclusion

We have examined a number of steadfast views about how one should epistemically behave in the face of peer disagreement. Despite their differences, all such views employ, what I have called, the standard strategy. Reflecting on these accounts, a certain pattern emerges which explains why they are susceptible to the same problems. It will also help illuminate the conciliationists' reactions to such views. Let me explain.

To show how epistemic peers can reasonably maintain their beliefs after full disclosure, the steadfast views, employing the standard strategy, tend to postulate some sort of epistemic asymmetry between the parties involved in a dispute. We have seen that different steadfast views interpret the asymmetry differently. Some construe it in terms of pragmatic considerations (Moffett) and some in terms of different epistemic norms (Goldman) while others appeal to different error theories (Bergmann). Thus, it seems that, to explain the rationality of the agents' first-order beliefs about a particular proposition p in the case of a disagreement, we have to go one level up and assume that the agents have certain fundamental beliefs about, say, which epistemic norms are correct and justification-conferring or which evidence or insights are genuine. For Goldman it was the postulation of the objectively justified beliefs about the correct E-norms that grounded incompatible but rational attitudes towards the target proposition while, for Bergmann, it was the agents' rational beliefs about the pertinent error theories that allowed them to rationally maintain their beliefs after full disclosure. Despite its clear relevance to the question of the possibility of reasonable disagreement, the standard strategy is apt to make the steadfast views that employ it vulnerable to the following general problems.

1. Although the postulation of certain antecedent rational beliefs, on the part of the epistemic peers, is, as we have just seen, necessary to explain how

[11] At the end of his discussion, Bergmann makes some remarks which seem to suggest that he is not willing to attach much epistemic significance to the distinction between possible and actual disagreement. First, he points out, correctly, that the fact that it is possible for things to seem to us perceptually just as they do even if we are the victims of an evil demon does not imply that we should question the reliability of our perceptual beliefs. Likewise, he claim, the fact that we think our peer is mistaken, despite being equally confident as we are in our beliefs, does not justify revising those beliefs. He realizes, however, that while his first case concerns the mere possibility of skepticism, the second case concerns an actual scenario in which we are being challenged. He thinks, however, that "the lesson is the same nevertheless" (Bergmann 2009, 545).

they can rationally disagree about a particular proposition, it makes the steadfast views susceptible to the charge of begging the question against conciliationism. For to assume that agents rationally hold such fundamental and incompatible beliefs is to assume that rational disagreement is possible. We saw that, on Goldman's account, agents are O-justified in believing incompatible E-norms to be correct while, for Bergmann, they enjoy rational confidence in different error theories. So both accounts try to show that rational disagreement is possible at the level of the target proposition only by assuming that it is possible at another (more fundamental) level.

2. Another problem for the standard strategy is whether it can establish the possibility of *both* reasonable disagreement and mutually recognized reasonable disagreement after full disclosure. For what seems necessary for the possibility of reasonable disagreement seems to render mutually recognized reasonable disagreement impossible. For example, while, for Bergmann, the assumption that agents hold different error theories is necessary in order to explain how they can remain rational after full disclosure, that very same assumption seems to hinder the possibility of *mutually recognized* peer disagreement. This is, of course, to be expected if an agent is to remain justified in believing her error theory after full disclosure. What is, at most, mutually recognizable by epistemic peers is that the other enjoys iterated justification in her belief about the target proposition. The same thing is true about Goldman's OBR-account which explicitly endorses only iterated justification for the target proposition. Regardless of whether such an attenuated notion is sufficient to establish the possible of reasonability disagreement, the problem with this suggestion is that one could already acknowledge its tenability without adverting to the various epistemological frameworks promulgated by the steadfast views such as Goldman's or Bergmann's. Iteratively justified belief, as non-culpable belief that is responsibly formed, is what is to be expected from agents who recognize each other as rational peers.

3. Finally, the implementation of the standard strategy is apt to lead to the violation of an adequacy condition on an acceptable theory of reasonable disagreement, namely, the recognition of the epistemological significance of the distinction between reasonable disagreement in isolation and reasonable disagreement after disclosure. We saw that once the standard strategy is used at the service of explaining the reasonability of disagreement it is bound to turn into a defense mechanism that would eventually obviate the epistemological significance of the aforementioned distinction.

For these reasons, I conclude that the steadfast views had better look into options other than adopting the standard strategy.

References

Alston, William P. 1988. "The Deontological Conception of Epistemic Justification." Reprinted in his *Epistemic Justification: Essays in the Theory of Knowledge*. Ithaca: Cornell University Press, 1989.

Bergmann, Michael. 2009, "Rational Disagreement after Full Disclosure." *Episteme: A Journal of Social Epistemology* 6, 3: 336-353.

Christensen, David. 2007. "Epistemology of Disagreement: the Good News." *The Philosophical Review* 116: 187–217.

Elga, Adam. 2007, "Reflection and Disagreement." *Noûs* 41, 3: 478-502.

Feldman, Richard. 2006. "Epistemological Puzzles about Disagreement." In *Epistemology Futures*, ed. Stephen Hetherington, 216-236. New York: Oxford University Press.

Feldman, Richard. 2007. "Reasonable Religious Disagreements." In *Philosophers without Gods: Meditations on Atheism and the Secular Life*, ed. Louise M. Antony, 194–214. Oxford: Oxford University Press.

Goldman, Alvin I. 1988. "Strong and Weak Justification." In *Philosophical Perspectives*, Volume 13, ed. James E. Tomberlin. Atascadero, CA: Ridgeview.

____. 2010. "Epistemic Relativism and Reasonable Disagreement." In *Disagreement*, eds. Richard Feldman and Ted Warfield. New York: Oxford University Press.

Harman, Gilbert. 1986. *Change in View: Principles of Reasoning*. Cambridge: The MIT Press.

Kelly, Thomas. 2005. "The Epistemic Significance of Disagreement." In *Oxford Studies in Epistemology*, vol. 1, eds. Tamar Szabo Gendler and John Hawthorne, 167–96. New York: Oxford University Press.

Moffett, Marc A. 2007. "Reasonable Disagreement and Rational Group Inquiry." *Episteme: A Journal of Social Epistemology* 4, 3: 352-367.

Sklar, Lawrence. 1975. "Methodological Conservatism." *Philosophical Review* LXXIV: 186-191.

Sosa, Ernest. 2010. "The Epistemology of Disagreement." In *Social Epistemology*, eds. Adrian Haddock, Alan Millar and Duncan Pritchard. Oxford: Oxford University Press.

Vahid, Hamid. 2004. "Varieties of Epistemic Conservatism." *Synthese* 141: 97-122.

Discussion Notes/Debate

For True Conditionalizers Weisberg's Paradox is a False Alarm

Franz Huber

Abstract: Weisberg (2009) introduces a phenomenon he terms *perceptual undermining*. He argues that it poses a problem for Jeffrey conditionalization (Jeffrey 1983), and Bayesian epistemology in general. This is *Weisberg's paradox*. Weisberg (2014) argues that perceptual undermining also poses a problem for ranking theory (Spohn 2012) and for Dempster-Shafer theory (Shafer 1976). In this note I argue that perceptual undermining does not pose a problem for any of these theories: for true conditionalizers Weisberg's paradox is a false alarm.

Keywords: perceptual undermining, Weisberg's paradox, Jeffrey conditionalization, Bayesian epistemology, ranking theory

1. Weisberg's Paradox

Weisberg's paradox consists in the inconsistency of four seemingly plausible constraints. It arises from the following example. Let D be the proposition that the sock really is red, and let F be the hypothesis that the lighting makes all socks look red. At time t_0 Sophia does not believe that the sock really is red and has a low degree of belief in D. Between t_0 and time t_1 she has a visual experience by looking at the sock. This visual experience between t_0 and t_1 causes her, among other things, to form the belief that the sock really is red at t_1; it leads to an increase in her degree of belief in D at t_1. At time t_2 she becomes certain that the lighting makes all socks look red; she assigns a maximal degree of belief to F at t_2. Since F is supposed to undermine the visual experience she has had between t_0 and t_1, this should make her drop her newly acquired belief that the sock really is red at t_2 again, and lower her degree of belief in D at t_2 back to what it was at t_0.

In a probabilistic setting this story is claimed to give rise to the following three constraints:

0. At time t_0 Sophia's degree of belief in D is low, say $\Pr_0(D) = .1$.

1a. At time t_1 Sophia's degree of belief in D is high, say $\Pr_1(D) = .9$.

2. At time t_2 Sophia's degree of belief in D is low again, $\Pr_2(D) = \Pr(D \mid F) = Pr_0(D) = .1$.

© Symposion, 1, 1 (2014): 111-119

Sophia's degree of belief function at time t_2, Pr_2, comes from her degree of belief function at time t_1, Pr_1, by an application of strict conditionalization to the hypothesis *F*. Strict conditionalization is the following update rule:

> **Update Rule 1 (Strict Conditionalization)** If $\mathrm{Pr}_0(\cdot)$: $\mathcal{A} \to \mathbb{R}$ is the ideal doxastic agent's probability measure at time t_0, and between t_0 and t_1 she becomes certain of the proposition $E \in \mathcal{A}$, $\mathrm{Pr}_0(E) > 0$, in the sense that $\mathrm{Pr}_1(E) = 1$, and she does not become certain of any logically stronger proposition (and her probabilities are not directly affected in any other way such as forgetting etc.), then her probability measure at time t_1 should be $\mathrm{Pr}_1(\cdot)$:$\mathcal{A} \to \mathbb{R}$
>
> $$\mathrm{Pr}_1(\cdot) = \mathrm{Pr}_E(\cdot) = \mathrm{Pr}_0(\cdot \mid E) = \mathrm{Pr}_0(\cdot \cap E) \,/\, \mathrm{Pr}_0(E)\,.$$

It is important to note that *E* is assumed to be the *total* evidence the ideal doxastic agent receives between time t_0 and time t_1.

Sophia's degree of belief function at time t_1, Pr_1, is assumed to come from her degree of belief function at time t_0, Pr_0, by an application of Jeffrey conditionalization to the evidential partition $\{D, \overline{D}\}$ with input parameters $\mathrm{Pr}_1(D) = .9$ and $\mathrm{Pr}_1(\overline{D}) = .1$. This assumption will turn out to be crucial. Hence it is stated as an independent constraint:

> 1b. Sophia's degree of belief function at time t_1, Pr_1, comes from her degree of belief function at time t_0, Pr_0, by an application of Jeffrey conditionalization to the evidential partition $\{D, \overline{D}\}$ with input parameters $\mathrm{Pr}_1(D) = .9$ and $\mathrm{Pr}_6(\overline{D}) = .1$.

Jeffrey conditionalization is the following update rule:

> **Update Rule 2 (Jeffrey Conditionalization, Jeffrey 1983)** *If* $\mathrm{Pr}_0(\cdot)$:$\mathcal{A} \to \mathbb{R}$ *is the ideal doxastic agent's probability measure at time* t_0*, and between* t_0 *and* t_1 *her probabilities on the evidential partition* $\{E_i \in \mathcal{A} : i \in I\}$ *directly change to* $p_i \in \mathbb{R}$*, where* $\sum_{i \in I} p_i = 1$*, and* $p = 0$ *if* $\mathrm{Pr}_0(E_i) = 0$*, and her positive probabilities do not directly change on any finer partition (and her probabilities are not directly affected in any other way such as forgetting etc.), then her probability measure at time* t_1 *should be* $\mathrm{Pr}_1(\cdot)$:$\mathcal{A} \to \mathbb{R}$
>
> $$\mathrm{Pr}_1(\cdot) = \mathrm{Pr}_{E_i \to p_i}(\cdot) = \sum_{i \in I} \mathrm{Pr}_0(\cdot \mid E_i) \cdot p_i.$$

It is important to note that the evidential partition $\{E_i \in \mathcal{A} : i \in I\}$ and the input parameters $p_i \in \mathbb{R}$ are assumed to be a *complete* description of all doxastically relevant events that happen between t_0 and t_1. In our example it is important to note that constraint (1b) amounts to the assumption that the only doxastically relevant effect of the visual experience between t_0 and t_1 is that Sophia's degrees of belief in *D* and $\overline{D}$ change to $\mathrm{Pr}_1(D) = .9$ and $\mathrm{Pr}_1(\overline{D}) = .1$. In

particular, it follows from constraint (1b) that F is not directly affected, in any doxastically relevant way, by Sophia's visual experience between t_0 and t_1. Among other things, this means that, at t_1, Sophia does not also form a belief about how she came to be more confident in D (or, if she has a belief at t_1 about how she came to be more confident in D, then it is the same belief that she had at t_0, *before* she had the visual experience). In order to simplify matters, let us assume, as is reasonable, that, at t_0, Sophia has no particular belief about what will happen between t_0 and t_1. Then constraint (1b) implies that, at t_1, Sophia remains agnostic as to whether it was by vision, or by some other form of perception, or by testimony, or by clairvoyance that she became more confident in D.

In other words, unless she already does so at t_0, Sophia does not believe at t_1 that the experience she undergoes between t_0 and t_1 is a *visual* experience. For all she believes at t_1, the experience she undergoes between t_0 and t_1 may not even be a *perceptual* experience. Indeed, given our reasonable assumption, for all she believes at t_1, what happens between t_0 and t_1 may not even be an *experience* of hers. *We* know this from the way the story was told, but, according to constraint (1b), *she* does not.[1]

So far, so good. Now the allegedly bad news. Jeffrey conditionalization is *rigid*. This implies that Jeffrey conditionalization preserves probabilistic independence of the members of the evidential partition. If a proposition A is independent of the evidential proposition D according to Sophia's degree of belief function at time t_0, $\Pr_0(A \mid D) = \Pr_0(A)$, then A is also independent of D according to her degree of belief function at a time t_1, $\Pr_1(A \mid D) = \Pr_1(A)$.

Weisberg (2009; 2014) thinks that this poses a problem for Jeffrey conditionalization, and Bayesian epistemology in general. He does so, because there is the following fourth constraint:

> 3. F is independent of D according to Sophia's degree of belief function at time t_0, $\Pr_0(F \mid D) = \Pr_0(F)$.

Weisberg (2009; 2014) is, of course, free to stipulate any constraints he wants. However, in order to evaluate the joint plausibility of his constraints, we need to understand what this condition says, and what it does not say. The condition says: at t_0, the degree to which Sophia believes that the lighting makes all socks look red has no bearing on the degree to which she believes that the

[1] If, contrary to what constraint (1b) implies, Sophia had even the slightest of hunches about what might have caused the changes in her degrees of belief in D and in $\overline{D}$, the evidential partition would include other propositions besides D and $\overline{D}$. In this case Weisberg's paradox would not arise.

sock really is red. The condition does *not* say: at t_0, whether or not the lighting makes all socks look red has no bearing on whether or not the sock really is red.

Together with the rigidity of Jeffrey conditionalization this fourth constraint implies that F is independent of D according to Sophia's degree of belief function at time t_1, $\mathrm{Pr}_1(F \mid D) = \mathrm{Pr}_1(F)$. However, if F is independent of D according to Sophia's degree of belief function at time t_1, $\mathrm{Pr}_1(F \mid D) = \mathrm{Pr}_1(F)$, then D is independent of F according to her degree of belief function at time t_1, $\mathrm{Pr}_1(D \mid F) = \mathrm{Pr}_1(D)$. In this case her degree of belief in D at time t_2 equals her degree of belief in D at time t_1, $Pr_2(D) = \mathrm{Pr}_1(D) = .9$. This contradicts the third constraint according to which $\mathrm{Pr}_2(D) = .1$.

Weisberg's paradox consists in this inconsistency. In the next section I will defend Jeffrey conditionalization and argue that Weisberg's paradox is resolved once we notice the implications of constraints (0-1) and (3) – in particular, (1b) – which make constraint (2) utterly implausible.

2. Jeffrey Conditionalization Defended

Weisberg's paradox consists in the inconsistency of the four seemingly plausible constraints (0-3), where Jeffrey conditionalization is part of constraint (1). For Weisberg (2009; 2014) the culprit is constraint (1) with Jeffrey conditionalization. I want to defend Jeffrey conditionalization.

Constraints (0-1) imply that what Sophia has experienced between t_0 and t_1 when looking at the sock results in an increase in her degree of belief in the proposition D that the sock really is red: $\mathrm{Pr}_1(D) > \mathrm{Pr}_0(D)$. However, in conjunction with the allegedly plausible constraint (3) constraints (0-1) also imply that what Sophia has experienced between t_0 and t_1 makes her hold onto her degree of belief in the hypothesis F that the lighting makes all socks look red: $\mathrm{Pr}_1(F) = \mathrm{Pr}_0(F)$. I want to use this consequence of constraints (0-1) and (3) to motivate my defense of Jeffrey conditionalization.

Roughly speaking, the former inequality $\mathrm{Pr}_1(D) > \mathrm{Pr}_0(D)$ says that at t_1 Sophia thinks that what she has experienced between t_0 and t_1 – or better, as Sophia may not even believe that it was an experience: what has happened between t_0 and t_1 – is related to D. The difference between $\mathrm{Pr}_1(D)$ and $\mathrm{Pr}_0(D)$, in whichever way it is measured, reflects how likely she thinks, at t_1, that what has happened between t_0 and t_1 is related to D. The latter equation $\mathrm{Pr}_1(F) = \mathrm{Pr}_0(F)$ says, roughly, that at t_1 she thinks that what has happened between t_0 and t_1 has nothing to do with F. The difference between $\mathrm{Pr}_1(F)$ and $\mathrm{Pr}_0(F)$, in whichever way it is measured, is nil, and that is how likely she thinks it, at t_1, that what has happened between t_0 and t_1 is related to F. In particular, this latter equation says,

roughly, that, at t_1, Sophia thinks that F is not a potential underminer for what has happened between t_0 and t_1.

Less roughly speaking, on the one hand constraints (0-1) say that Sophia's belief in D is directly affected by what happens between t_0 and t_1 in such a way that she ends up being more confident in D at t_1 than she was at t_0. On the other hand constraints (0-1) say that Sophia's belief in F is *not* directly affected by what happens between t_0 and t_1. Constraints (0-1) and the allegedly plausible stipulation (3) add to this that Sophia's belief in F is not indirectly affected by what happens between t_0 and t_1 either. Together constraints (0-1) and (3) thus say that Sophia's belief in F is neither directly nor indirectly affected by what happens between t_0 and t_1. In particular, constraints (0-1) and (3) say that, at t_1, Sophia thinks that F is not a potential underminer for what has happened between t_0 and t_1.[2]

In other words, constraints (0-1) and (3) imply that at t_1 Sophia thinks that what has happened between t_0 and t_1 may be due the fact that the sock really is red, D, but is definitely not due to the fact that the lighting makes all socks look red, F. Given this consequence of constraints (0-1) and (3), constraint (2) clearly should be rejected: constraint (2) says that, at t_1, Sophia thinks that F is a potential underminer for what has happened between t_0 and t_1, whereas it follows from constraints (0-1) and (3) that, at t_1, she thinks that F is not a potential underminer for what has happened between t_0 and t_1.

Here is a different way of putting things. Suppose we modeled the change of Sophia's degree of belief function from t_0 to t_1 by an application of strict conditionalization to the appearance proposition D^* that the sock appears to be red, $\Pr_1(\cdot) = \Pr_0(\cdot \mid D^*)$. In this case Sophia would believe that the apparent color of the sock is relevant to the actual color of the sock, $\Pr_1(D) = \Pr_0(D \mid D^*) > \Pr_0(D)$. More importantly, however, she would presumably also believe that the apparent color of the sock is relevant to the hypothesis that the lighting makes all socks look red, $\Pr_1(F) = \Pr_0(F \mid D^*) > \Pr_0(F)$. Consequently we would not have the consequence $\Pr_1(F) = \Pr_0(F \mid D^*) = \Pr_0(F)$.

[2] The conjunction of constraints (0-1) and (3) is logically strictly stronger than its consequence $\Pr_0(F) = \Pr_1(F)$. The latter equation can be true if constraints (0-1) or (3) are false, because the various effects on Sophia's belief in F may "cancel out." For instance, what happens between t_0 and t_1 may affect Sophia's belief in the four propositions $X \cap F$, $X \cap \overline{F}$, $\overline{X} \cap F$, $\overline{X} \cap \overline{F}$,for some proposition X, even though Sophia's degrees of belief in $X \cap F$ and $\overline{X} \cap F$ may sum to the same number at t_0 and at t_1. It is the conjunction of constraints (0-1) and (3), not its consequence, that says that, at t_1, Sophia thinks that F is not a potential underminer for what has happened between t_0 and t_1.

Now in Weisberg's example there is no appearance proposition D^* that Sophia becomes certain of between t_0 and t_1. Instead of an application of strict conditionalization to the appearance proposition D^* we have an application of Jeffrey conditionalization to the evidential partition $\{D, \overline{D}\}$ that is caused by some visual experience d^*. Instead of the appearance proposition D^* that Sophia becomes certain of, the doxastically relevant effects of the visual experience d^* are now described more indirectly: they correspond to the differences between Sophia's degrees of belief in D and $\overline{D}$ at t_0 and her degrees of belief in D and $\overline{D}$ at t_1. The visual experience d^* is non-propositional evidence that is reflected in the different shapes of Sophia's two degree of belief functions at t_0 and at t_1 on the evidential partition $\{D, \overline{D}\}$. Most importantly, in contrast to the above case where Sophia becomes certain of an *appearance* proposition, and thus learns something about what drives the change in her degrees of belief, constraint (1b) *ex*cludes information about what drives the change in Sophia's degrees of belief in D and $\overline{D}$. For all Sophia believes, the change in her degrees of belief in D and $\overline{D}$ may be due to a lapse of rationality.

Given these consequences it becomes clear that, in the presence of constraints (0-1) and (3) – in particular, constraint (1b) – constraint (2) is utterly implausible. If Sophia does not believe that the change in her degrees of belief in D and $\overline{D}$ between t_0 and t_1 has been caused by her vision, why should she believe that F can undo this doxastically relevant effect of what has happened between t_0 an t_1? The way the story is told makes it clear to *us* that it was her vision that caused her increase in confidence in D. However, as long as this information is not also made available to *her*,[3] there is no reason whatsoever to assume that her becoming certain of F should have any effect at all on her degree of belief in D.

To be sure, there are potential underminers G for Sophia's visual experience that takes place between t_0 and t_1, and constraint (2) can be satisfied. However, if constraint (2) is satisfied for some potential underminer G, then this potential underminer G must violate constraint (3) or constraint (1). Constraint (3) is violated by G if G is not independent of D according to $\Pr_0$. Constraint (1) is violated by G if Jeffrey conditionalization is applied to an evidential partition $\{E_i: i \in I\}$ whose members are not all logically independent of the potential

[3] One way of making this information available to Sophia is by using Shafer (1985)'s notion of a *protocol*. Halpern (2003, ch. 6) uses protocols to solve Freund (1965)'s puzzle. Spohn (2012, sct. 9.3) uses protocols to solve the puzzle of the three prisoners (Mosteller 1965, problem 13) that is also known as the Monty Hall problem. Halpern (ms) uses protocols to solve the Doomsday Argument and the Sleeping Beauty Problem. A different proposal is discussed in the Appendix.

underminer G, and which must thus be more fine-grained than $\{D, \overline{D}\}$, say, $\{D \cap F, \overline{D} \cap F, D \cap \overline{F}, \overline{D} \cap \overline{F}\}$.

If constraint (3) is violated by G, but constraint (1) holds, then G is indirectly affected by the visual experience that takes place between t_0 and t_1, and Jeffrey conditionalization governs this indirect way of being affected. In this case G figures in the output of Jeffrey conditionalization, which tells Sophia, among other things, what to believe about G. If constraint (1) is violated by G, then G is directly affected by the visual experience that takes place between t_0 and t_1, and Jeffrey conditionalization has to be applied to an evidential partition some of whose members are logically dependent on G. In this case G figures in the input of Jeffrey conditionalization, and Sophia has to specify her new degrees of belief for all these members, including those that logically depend on G, before Jeffrey conditionalization can be applied.

Either way there is no problem for Jeffrey conditionalization. We just have to realize that, while a change of Sophia's degrees of belief from t_0 to t_1 can be undermined, it can only be undermined by a hypothesis G that is doxastically affected by what drives the former change. This can happen indirectly by G not being probabilistically independent of the members of the evidential partition and thus violating the allegedly plausible stipulation (3). Or it can happen directly by G not being logically independent of all members of the evidential partition to which Jeffrey conditionalization is applied. Which is determined by experience, not by methodology.

For these reasons I conclude that Weisberg's paradox may affect the applicability of Jeffrey conditionalization, but not its validity. Weisberg's paradox does not undermine Jeffrey's rule of conditionalization: for true conditionalizers Weisberg's paradox is a false alarm. Parallel considerations show that perceptual undermining does not pose a problem for ranking theory or Dempster-Shafer theory either.

3. Appendix

One way of making the information that it was her vision that caused her increase in confidence in D available to Sophia is by using Shafer (1985)'s notion of a *protocol*. Protocols have proved to be a powerful tool in solving paradoxes. For instance, Halpern (2003, ch. 6) uses protocols to solve Freund (1965)'s puzzle. Spohn (2012, sct. 9.3) uses protocols to solve the puzzle of the three prisoners (Mosteller 1965, problem 13) that is also known as the Monty Hall problem. Halpern (ms) uses protocols to solve the Doomsday Argument and the Sleeping Beauty Problem. It is only natural that protocols can also be used to solve Weisberg's Paradox.

A different proposal for making the information that it was Sophia's vision that caused her increase in confidence in D available to Sophia is presented in Gallow (2014). Gallow (2014) first proposes the following update rule:

> **Update Rule 3 (Gallow Conditionalization I, Gallow 2014)** If $\mathrm{Pr}_0(\cdot)$:$\mathcal{A} \to \mathbb{R}$ is the ideal doxastic agent's probability measure at time t_0, and between t_0 and t_1 she receives total evidence of the form $\{(T_i, E_i): i \in I\}$, where the $T_i \in \mathcal{A}$ form a partition of the underlying set of possible worlds W, and her probabilities are not directly affected in any other way such as forgetting etc., then her probability measure at time t_1 should be $\mathrm{Pr}_1(\cdot)$:$\mathcal{A} \to \mathbb{R}$,
>
> $$\mathrm{Pr}_1(\cdot) = \sum_{i \in I} \mathrm{Pr}_0(\cdot \mid T_i \cap E_i) \cdot \mathrm{Pr}_0(T_i).$$

The interpretation of a pair (T_i, E_i) is that the ideal doxastic agent's evidence is E_i, provided T_i is the case. For instance, Sophia's total evidence may be that the sock really is red, provided the lighting does not make all socks look red and everything else is normal as well; and nothing otherwise: $\{(N, E), (\neg N, W)\}$.

Gallow conditionalization I is an instance of Jeffrey conditionalization on the evidential partition $\{T_i \cap E_i, T_i \cap \overline{E_i}: i \in I\}$ with input parameters $\mathrm{Pr}_1(T_i \cap E_i) = \mathrm{Pr}_0(T)$ and $\mathrm{Pr}_1(T_i \cap \overline{E_i}) = 0$. Unfortunately it does not allow the ideal doxastic agent to ever change her confidence in any of the theories T_i, i.e. $\mathrm{Pr}_1(T_i) = \mathrm{Pr}_0(T_i)$.

For this reason Gallow (2014) generalizes his first update rule to the following second update rule:

> **Update Rule 4 (Gallow Conditionalization II, Gallow 2014)** *If* $\mathrm{Pr}_0(\cdot)$:$\mathcal{A} \to \mathbb{R}$ *is the ideal doxastic agent's probability measure at time* t_0*, and between* t_0 *and* t_1 *she receives total evidence of the form* $\{(T_i, E_i): i \in I\}$*, where the* $T_i \in \mathcal{A}$ *form a partition of the underlying set of possible worlds W, and her probabilities are not directly affected in any other way such as forgetting etc., then her probability measure at time* t_1 *should be* $\mathrm{Pr}_1(\cdot)$:$\mathcal{A} \to \mathbb{R}$,
>
> $$\mathrm{Pr}_1(\cdot) = \sum_{i \in I} \mathrm{Pr}_0(\cdot \mid T_i \cap E_i) \cdot \mathrm{Pr}_0(T_i) \cdot \Delta_i,$$
>
> *where* Δ_i *is a non-negative real number representing the degree to which the total evidence* $\{(T_i, E_i): i \in I\}$ *(dis)confirms theory* T_i.

Contrary to what Gallow (2014, 21) claims his second update rule does not generalize Jeffrey conditionalization. Indeed, quite the opposite is the case. Gallow conditionalization II is also just an instance of Jeffrey conditionaliaztion on the evidential partition $\{T_i \cap E_i, T_i \cap \overline{E_i}: i \in I\}$ with input parameters $\mathrm{Pr}_1(T_i \cap E_i) = \mathrm{Pr}_0(T_i) \cdot \Delta_i$ and $\mathrm{Pr}_1(T_i \cap \overline{E_i}) = 0$. (As Gallow (2014, 25ff) shows, the sum

$\sum_{i\in I} \mathbf{Pr_0}(T_i) \cdot \Delta_i = \sum_{i\in I} \mathbf{Pr_1}(T_i \cap E_i) + \mathbf{Pr_1}(T_i \cap \overline{E_i})$ equals 1, and so the constraints of Jeffrey conditionalization are satisfied.)[4]

References

Freund, John E. 1965. "Puzzle or paradox?" *American Statistician* 19: 29-44.

Gallow, Dmitri J. 2014. "How to Learn from Theory-Dependent Evidence; or Commmutativity and Holism: A Solution for Conditionalizers." *British Journal for the Philosophy of Science*.

Halpern, Joseph Y. 2003. *Reasoning About Uncertainty*. Cambridge, Massachusetts: MIT Press.

____. 2014. *The Role of the Protocol in Anthropic Reasoning*. Unpublished manuscript. https://www.cs.cornell.edu/home/halpern/papers/doomsday.pdf (accessed July 7, 2014).

Jeffrey, Richard C. 1983. *The Logic of Decision*, 2nd ed. Chicago: University of Chicago Press.

Mosteller, Frederick. 1965. *Fifty Challenging Problems in Probability with Solutions*. Reading: Addison-Wesley.

Shafer, Glenn. 1976. *A Mathematical Theory of Evidence*. Princeton: Princeton University Press.

____. 1985. "Conditional probability." *International Statistical Review* 53: 261-277.

Spohn, Wolfgang. 2012. *The Laws of Belief. Ranking Theory and its Philosophical Applications*. Oxford: Oxford University Press.

Weisberg, Jonathan. 2009. "Commutativity or Holism? A Dilemma for Conditionalizers." *British Journal for the Philosophy of Science* 60: 793-812.

____. 2014. "Updating, Undermining, and Independence." *British Journal for the Philosophy of Science*.

[4] I am grateful to Kevin Kuhl, Wolfgang Spohn, and Jonathan Weisberg for comments on earlier versions of this note. My research was supported by the Canadian SSHRC through its Insight program and by the Connaught Foundation through its New Researcher program.

Information about authors

Scott F. Aikin is Assistant Professor of Philosophy at Vanderbilt University. He works primarily in epistemology and ancient philosophy. He is the author of *Epistemology and the Regress Problem* (Routledge 2011) and *Evidentialism and the Will to Believe* (Bloomsbury 2014), and the co-author (with Robert B. Talisse) of *Why We Argue (And How We Should)* (Routledge, 2014), *Reasonable Atheism* (Prometheus Books, 2011), and *Pragmatism: A Guide for the Perplexed* (Continuum Books, 2008). He edited *The Regress Problem: Meta-theory, Development, and Criticism* (Special issue of *Metaphilosophy*, 2014, with Jeanne Peijnenburg), *The Pragmatism Reader: From Peirce to the Present* (with Robert B. Talisse, Princeton University Press, 2011), and *Thinking about Logic: Classic Essays* (with Steven M. Cahn and Robert B. Talisse, Westview Press 2010). Contact: scott.f.aikin@vanderbilt.edu.

Ionuţ-Alexandru Bârliba is Post-PhD Fellow in the SOP HRD/159/1.5/133675 Project, Romanian Academy, Iaşi Branch. He recently became assistant researcher at the "Gh. Zane" Institute of Economic and Social Research of the Romanian Academy, Iaşi Branch. He holds a PhD from "Alexandru Ioan Cuza" University, Iaşi with a thesis on Søren Kierkegaard's philosophy (defended in 2012). His main areas of interest are metaphysics, hermeneutics, existentialism and Romanian philosophy. His recent papers include: "The Possibility of a Happy Fall" (*Meta*, 2013) and "The Pressure of Leveling. The Single Individual vs. the Crowd at Søren Kierkegaard" ("Alexandru Ioan Cuza" University, forthcoming, in Romanian). E-mail: ionutbarliba@gmail.com.

Franz Huber is an assistant professor in the Department of Philosophy at the University of Toronto. His research interests are in formal epistemology, general philosophy of science, and philosophical logic. With Jonathan Weisberg and many others he is editing *Ergo, An Open Access Journal of Philosophy*. His most recent articles include: "What is the Permissibility Solution a Solution of? – A Question for Kroedel" (*Logos & Episteme*, forthcoming), "What Should I Believe About What Would Have Been the Case?" (*Journal of Philosophical Logic*, 2014), and "New Foundations for Counterfactuals" (*Synthese*, 2014). Contact: franz.huber@utoronto.ca.

Dale Jacquette is ordentlicher Professor für Philosophie, Abteilung Logik und theoretische Philosophie (Senior Professorial Chair in Logic and Theoretical Philosophy), at Universität Bern, Switzerland. He is the author of numerous books and articles in topics and fields such as intentionality, logic, metaphysics, ethics, philosophy of mind, epistemology, philosophy of science, aesthetics or history of philosophy. His books include *Logic and How it Gets That Way*

(Acumen Books, 2010), *Philosophy of Mind: The Metaphysics of Consciousness* (Continuum Books, 2009), *Dialogues on the Ethics of Capital Punishment* (Rowman & Littlefield, 2009), *Journalistic Ethics: Moral Responsibility in the Media* (Prentice Hall, 2007) *The Philosophy of Schopenhauer* (Acumen Books, 2005), and *Ontology* (Acumen Books, 2002). Some of his edited books are *Russell versus Meinong: The Legacy of 'On Denoting'* (with Nicholas Griffin, Routledge, 2009), *Philosophy of Logic* (Elsevier, 2007), *The Cambridge Companion to Brentano* (Cambridge University Press, 2004), *Philosophy of Mathematics: An Anthology* (Blackwell, 2002), *Philosophy of Logic: An Anthology* (Blackwell, 2002), and *A Companion to Philosophical Logic* (Blackwell, 2002). Contact: dale.jacquette@philo.unibe.ch.

Franck Lihoreau is a Research Fellow at the University of Coimbra, Portugal.. His main areas of research include epistemology, philosophy of language, philosophical logic, and metaphysics. He has publications in all of these and is the editor and co-editor of several volumes, including *Knowledge and Questions* (Rodopi, 2008), *Truth in Fiction* (Ontos Verlag, 2011), *Epistemology, Context, and Formalism* (Springer, 2014), and *Dialogue, Rationality, and Formalism* (Springer, 2014). Contact: franck.lihoreau@gmail.com.

Clayton Littlejohn is Lecturer with the Department of Philosophy at King's College London. His current research focuses primarily on knowledge-first epistemology. He is the author of *Justification and the Truth-Connection* (Cambridge University Press, 2012) and *This is Epistemology* (Wiley-Blackwell, forthcoming), an introduction to epistemology. Contact: cmlittlejohn@gmail.com.

Emily McGill-Rutherford is a graduate student at Vanderbilt University. She works in social and political philosophy and feminist philosophy, and is currently completing a dissertation titled *What Women Want: Feminism, Liberalism, and Relational Autonomy*. Contact: emily.mcgill@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 most recent articles include: "Varieties of Pragmatic Encroachment" (*Acta Analytica*, 2014) "Skepticism, A Priori Skepticism and the Possibility of Error" (*International Journal for the Study of Skepticism*, 2013), and "Skepticism and the Liberal/Conservative Conceptions of Perceptual Justification" (*International Journal for the Study of Skepticism*, 2013). Contact: vahid@ipm.ir.

About the Journal

Symposion was published for the first time in 2003, as *Symposion – Revistă de ştiinţe socio-umane (Symposion – A Journal of Humanities)*, with the purpose of providing a supportive space for academic communication, dialog, and debate in various research fields in humanities and social sciences. *Symposion: Theoretical and Applied Inquiries in Philosophy and Social Sciences* shares the same purpose. Its main aim is to promote and provide open access to peer-reviewed, high quality contributions for scholars working in (analytic, continental or eastern) philosophy, other humanities disciplines, or social and political theory (both empirical and normative).

The old journal published mainly Romanian papers. The new *Symposion* is intented to be an international journal, welcoming contributions from around the world written in English, French or German, and only exceptionally in Romanian.

Symposion is open for original and innovative contributions in all humanities and social sciences, and will publish with priority papers (articles, discussion notes, review essays or book reviews) in the following research fields: aesthetics and philosophy of art, anthropology, history, and cultural studies, communication and media studies, economics and philosophy of economics, education and social work, epistemology and philosophy of science, metaethics, normative and applied ethics, hermeneutics and rhetoric, history of philosophy, law and philosophy of law, linguistics and philosophy of language, logic, philosophical logic, and philosophy of logic, metaphilosophy, metaphysics, philosophy of mind and cognitive science, philosophy and sociology of religion, psychology and metapsychology, social and political philosophy, social and political sciences. The journal is also available for scholars developing cross-disciplinary, multidisciplinary, interdisciplinary, or transdisciplinary research.

Symposion will be a quarterly journal, appearing at the end of January, April, July, and October.

Author Guidelines

1. Accepted Submissions

The journal accepts for publication articles, discussion notes, review essays and book reviews submitted exclusively to *Symposion* and not published, in whole or substantial part, elsewhere. Submission to *Symposion* is taken to imply that the same manuscript is not under consideration by another journal. The editors of *Symposion* 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: symposion.journal@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 and review essays must be no longer than 6000 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. Authors should use the author-date system for text citations and chicago style format for reference lists, as it is presented in Chicago Manual of Style.For details, please visit http://library.williams.edu/citing/styles/chicago2.php.

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. Papers in Romanian will be published only exceptionally. 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 last page of the paper.

8. Review Process

Symposion publishes standard submissions and invited papers. With the exception of invited contributions, all articles will be subject to a 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. Articles accepted will be published, as far as possible, in the order in which they are received and will appear in the journal under one of the following main sections: 1) Philosophical Reflections; 2) Explorations in Humanities; 3) Inquiries in Political Theory; 4) Social Science Investigations; and 5) Crossdisciplinary/Multidiciplinary/Interdisciplinary/Transdisciplinary Research. The authors should inform the editorial board about their option regarding the most appropriate section for their article. Inside each section, papers will be published mainly 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 working 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: Symposion Journal, 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 *Symposion* 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.

14. Permissions