Volume XII ◇ Issue 3

2021

Logos & Episteme

an international journal of epistemology

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

Contact address:

Institutul de Cercetări
Economice şi Sociale „Gh.Zane”

Iaşi, str.T.Codrescu, nr.2, cod 700481

Tel/Fax: 004 0332 408922

Email: logosandepisteme@yahoo.com

http://logos-and-episteme.acadiasi.ro/

https://www.pdcnet.org/pdc/bvdb.nsf/journal?openform&journal=pdc_logos-episteme

ISSN 2069-053

TABLE OF CONTENTS

RESEARCH ARTICLES

MAKE INFORMATION IN SCIENCE MEANINGFUL AGAIN

Javier ANTA

ABSTRACT: Although the everyday notion of information has clear semantic properties, the all-pervasive technical concept of Shannon information was defended being a non-semantic concept. In this paper I will show how this measure of information was implicitly 'semantized' in the early 1950s by many authors, such as Rothstein's or Brillouin's, in order to explain the knowledge dynamics underlying certain scientific practices such as measurement. On the other hand, I will argue that the main attempts in the literature to develop a quantitative measure of semantic information to clarify science and scientific measurements, such as Carnap-Bar-Hillel, or Dretske, will not successfully achieve this philosophical aim for several reasons. Finally, I will defend the use of a qualitative notion of semantic information within the information-theoretical framework MacKay to assess the informational dynamics underlying scientific practices, particularly measurements in statistical mechanics.

KEYWORDS: Shannon information, communication theory, measurement, semantic information

1. Introduction

It is often said that the task of a scientist is to obtain information about his or her field of research. But what do we mean here by 'information'? As the information philosopher Floridi[1] argues, the term 'information' encompasses a huge plurality of different concepts and different meanings.[2] One of these is the ordinary sense of information, of an indisputable semantic character (highlighted in his 'aboutness,' i.e., John has information *about* what happened) and linked to the ability to provide knowledge to someone. However, Claude Shannon's famous theory of communication[3] placed his quantitative concept of information at the center of the imaginary of the technical-scientific community since the 1950s, whose importance will expand until today. Interestingly, this author employed a measure-concept of

[1] See Luciano Floridi, *The Philosophy of Information* (Oxford University Press, 2011)

[2] See Olimpia Lombardi, Federico Holik, and Leonardo Vanni, "What is Shannon information?" *Synthese* 193, 7 (2016):1983-2012.

[3] See Claude E. Shannon, "A Mathematical Theory of Communication," in *Collected Papers*, eds. N. J. A. Sloane and A. D. Wyner (New York: IEEE Press, 1948 [1993]).

information from a certain sequence of symbols that was completely independent of the meaning of these symbols, which contradicted the semantic character of the ordinary meaning of this term.

The enormous intellectual impact that Shannon's theoretical proposal had prompted several authors to use the concept of theoretical-communicative information to understand how scientists acquire information about their objective phenomena. For example, Jerome Rothstein in 1951 systematically compared scientific measurement processes with a communication process between observer and observed system.[4] However, as I will argue in this paper, this kind of application of Shannon information would implicitly involve attempts to semantize (in a misleading fashion) the everyday meaning of this term. In this paper I will argue that the main attempts in the literature to employ a theoretical-communicative concept of information (or other alternatives also statistical-quantitative) to illuminate scientific practices will not succeed in carrying out this task in a robust way, even if conceptual alternatives to Carnap and Bar-Hillel[5] are developed or formally complemented to Dretske[6] in a way that is somehow sensitive to the semantic content of certain information elements. On the other hand, I will propose not a new quantitative measure of information adapted to these scientific scenarios, but rather to develop an informational interpretation (using certain elements developed by MacKay[7]) of the very representational tools used in scientific practices.

The plan for this paper is the following. Next, I will present the basis of Shannon's communication theory,[8] where a non-semantic concept of information was developed in order to statistically evaluate the transmission of signals. In Section 3, I will analyze how, despite these characteristics, Shannon's informational concept began to be used during the 1950s as an informational measure used to clarify certain scientific practices such as measurement,[9] implicitly providing it with a semantic character. Later on, I will detail the main quantitative-statistical theories of information developed under the intellectual impact (either as an alternative or as a complement) of Shannon's proposal. In Section 5 I will argue that this first and last would fail to give a satisfactory account for various reasons of the informational

[4] Jerome Rothstein, "Information, Measurement, and Quantum Mechanics," *Science* 114 (1951): 171-175.

[5] Rudolf Carnap and Yehoshua Bar-Hillel, "An Outline of a Theory of Semantic Information," *Technical Report* 247 (1952), Research Laboratory of Electronics, MIT.

[6] Fred Dretske, *Knowledge and the Flow of Information* (Cambridge: MIT Press, 1981).

[7] Donald MacKay, *Information, Mechanism, and Meaning* (Cambridge: MIT Press, 1969).

[8] Shannon, "A Mathematical Theory."

[9] See Rothstein, "Information, Measurement" and Leon Brillouin, *Science and Information Theory* (New York: Academic Press, 1956).

dynamics of scientific practice. Finally, I will develop an alternative informational proposal from MacKay's framework based on the definition of a qualitative concept of information, in order to philosophically assess certain scientific measurement practices.

2. Shannon's Communication Theory

In "A Mathematical Theory of Information," Claude Shannon first set out the foundations of his theoretical proposal regarding the statistical analysis of the transmission of continuous or discrete messages (i.e., sequences of symbols belonging to a set of symbols or alphabet) within certain communicative contexts.[10] These contexts are made up of (i) a 'source' that generates the message, (ii) a 'transmitter' that transforms the message into signals to be transmitted, (iii) a 'receiver' that reconstructs the message at the point of destination, and (iv) a 'communicative channel' as a means of transmitting messages. In not so general terms, his proposal is based fundamentally on developing an H function (this is called 'source entropy') that measures the average amount of information generated by the source through the transmission of a message si...sj, where the probability of occurrence of each particular symbol is determined by the probability distribution over the source:

$$H(S) = -\sum p(s_i) \log p(s_i) \quad (1)$$

Roughly speaking, this amount of H entropy measures the degree of improbability (also often interpreted in epistemic terms as 'degree of unexpectedness' or 'unpredictability') with which a particular sequence of symbols s_i occurs, reaching its maximum value when the probability distribution over the source is uniform. For example, the message 'XZV' will have a significantly higher entropy than the message 'SKY' if we consider the frequency of occurrence of the English letters (i.e. H(XZV) > H(SKY)), precisely because the simultaneous occurrence of the 'XZV' symbols in a message would be extremely uncommon. Thus, this measure of information cannot be defined only for a particular message (i.e. 'SKY'), but for a message with respect to the probability distribution defined for the source or the set of all possible source symbol sequences. It is precisely the so-called 'noisy coding theorem' that mathematically determines that the most optimal way to encode messages is by sequences of units with binary values (or 'bits') of 0 and 1.

One of the main characteristics of Shannon entropy is its disregard for the semantic content of the message: "Frequently, the messages have meaning; that is they refer to or are correlated according to some with certain physical or conceptual

[10] Shannon, "A Mathematical Theory."

entities. These semantic aspects of communication are irrelevant to the engineering problem."[11] In other words, it is completely irrelevant whether the sequence 'SKY' actually refers to the sky or to a particular brick in the street so that the amount of information measured by means of the Shannon H(SKY) entropy is, for example, 0.05. This lack of semantic sensitivity of Shannon measure[12] of the information of a message to its (plausible) conventional meaning was the subject of enormous controversy since the popularization of this theory in the 1950s until today,[13] mainly because of its radical difference with the usual sense of the term 'information,' synonymous with 'knowledge' and inseparably linked to semantic-epistemic and intentional properties,[14] i.e. information of A *about* X. Another important point to disregard Shannon entropy as a measure of meaning or semantic content of messages (or other structures) would be the following: the function H does not measure the amount of (not semantics, as we have just pointed out) information conveyed in the transmission of particular messages, but the average amount of information of a statistical assembly of possible messages.[15]

Another important aspect of this entropy measure H is its formal similarity (based on the use of probability distributions and the use of the logarithmic function for its pragmatic virtues) with the Boltzmann and Gibbs measure of statistical mechanical entropy "the form of H will be recognized as that of entropy as defined in certain formulations of statistical mechanics."[16] Despite this similarity, the latter measure (roughly) the probability that the exact microscopic state of the physical system is in a cell or region of the space of possible molecular values or phase space. This leads us to consider the choice of this name 'entropy' originally belonging to the field of statistical physics to name a measure of quantity of information. Tribes

[11] Shannon, "A Mathematical Theory," 3.

[12] Recently, M. Alistair ("The Semantics Latent in Shannon Information," *British Journal for the Philosophy of Science* 70, 1 (2019): 103-125) has argued that it would be possible to extract certain semantic properties from the statistical correlations modeled from Shannon's theoretical-communicational formalism. However, his argument depends on a proto-information theory developed by Turing in the 1940s, and not directly on Shannon's theory.

[13] See Ronald Kline, *The Cybernetics Moment: Or Why We Call Our Age the Age of Information* (Baltimore: Johns Hopkins University Press, 2015), Section 2.

[14] See Floridi, *Philosophy of Information*.

[15] See Lombardi, Holik and Vanni, "What is Shannon." Note that it is misleading to (although often convenient) say that one or the other message conveys unit information. The concept of information applies not to the individual messages (as the concept of meaning would), but rather to the situation as a whole, the unit information indicating that in this situation one has an amount of freedom of choice, in selecting a message, which it is convenient to regard as a standard or unit amount (Kline, *The Cybernetics Moment*, 132).

[16] Shannon, "A Mathematical Theory," 12.

and McIving reported (from an interview of Shannon) that it was von Neumann who suggested the name to exploit his deep misunderstanding within the scientific community.[17]

As Shannon himself recognized, his proposal is presented as a highly sophisticated development that builds on the analyses developed in the papers by Nyquist and Hartley (where the latter's measure of 'information' is mathematically identical with Shannon's entropy in the case of a uniform probability distribution over the symbols) during the 1920s, deeply forgotten beyond the walls of the Bell Laboratories. Unlike these pioneers, Shannon developed throughout the 1940s (culminating in his 1948 paper) an extensive technical proposal on how to statistically evaluate and optimize the transmission of discrete/continuous messages in both noisy and noiseless channels, modeling this communicative process as a Markov chain. In short, we must emphasize once again that Shannon information is intrinsically independent of the meaning and physical character of the informational elements: "Shannon's theory is a quantitative theory whose elements have no semantic dimension (...) Moreover, Shannon's theory is not tied to a particular physical theory, but is independent of its physical interpretation."[18]

The director of the Division of Natural Sciences at the Rockefeller Foundation, Warren Weaver, immediately appreciated the prospects of Shannon's theory, not only within the field of communication but also in other scientific domains. His role in the enormous immediate impact that communication theory had within the scientific community was pivotal, popularizing Shannon's excessively technical proposal for the general public (note that even engineers had difficulties in understanding his theses[19]) through an introductory commentary in the reprint of the original article in the famous (Shannon and Weaver) 1949 book.[20] Moreover, one of his main aims (more or less implicit) was to elevate Shannon's intellectual work to the Olympus of American science in which the physicist J. W.

---

[17] "[Shannon said:] 'My greatest concern was what to call it. I thought of calling it 'information', but the word was overly used, so I decided to call it 'uncertainty.' When I discussed it with John von Neumann, he had a better idea. Von Neumann told me, 'You should call it entropy, for two reasons. In the first place you uncertainty function has been used in statistical mechanics under that name. In the second place, and more importantly, no one knows what entropy really is, so in a debate you will always have the advantage" (Myron Tribus and Edward McIrvine, "Energy and information," *Scientific American* 225 (1971): 179-188).

[18] See Lombardi, Holik, and Vanni. "What is Shannon," 2000.

[19] Kline, *The Cybernetics Moment*, Chapter 2.

[20] Claude E. Shannon and Warren Weaver, *The Mathematical Theory of Communication* (Urbana: University of Illinois Press, 1949).

Gibbs was, but to generate a historical narrative in which Shannon's theory was the culprit of the physical domain statistical mechanics as developed by Boltzmann:

> Dr. Shannon's work roots back, as von Neumann has pointed out, to Boltzmann's observation, in some of his work on statistical physics (1984) that entropy is related to 'missing information,' inasmuch as it is related to the number of alternatives which remain possible to a physical system after all the macroscopically observable information concerning it has been recorded.[21]

In this direction, Weaver intended to extend the domain of application and popularization of Shannon's theoretical proposal from the narrow technical field of signal transmission to the transdisciplinary field of statistical thermal physics. However, in order for the concept of Shannon's theoretical-communicative entropy to have relevance within this field of physics along the lines suggested by von Neumann (i.e. information from an observer about an observed phenomenon), it would have to appeal indispensably to a notion of an evident semantic character, which would contradict the Shannonian dogma of "semantic irrelevance." It was Weaver himself who opened the door to the possibility of developing a theoretical account of strictly semantic information based on Shannon's proposal "But this does not mean that the engineering aspects are necessarily irrelevant to the semantic aspects."[22]

3. Naturalizing Information in Scientific Practices: Beyond the Non-Semantic Dogma

Just after Shannon's paper on communication theory was reissued in the famous book, it became extremely popular within the scientific community of that time. This popularity was accompanied by an enormous interest in the application of his technical proposal in disciplines highly disconnected from the transmission of signals, mainly molecular biology and thermal statistical physics. In the latter case, the choice of the name 'entropy' and the suggestive comments of von Neumann and Weaver played an indispensable role in the progressive informationalization of thermal physics during the 50s. This growing trend encouraged the belief that Shannon's theory would necessarily play a central role in understanding the knowledge-formation process underlying certain scientific practices.

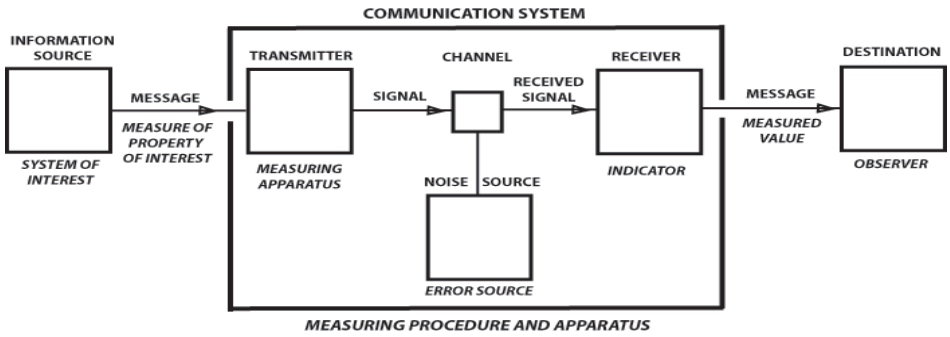
One of the most illustrative examples of this intellectual movement was the paper "Information, Measurement, and Quantum Mechanics" by Jerome Rothstein,

[21] See Shannon and Weaver, *The Mathematical Theory*, note 1.

[22] See Shannon and Weaver, *The Mathematical Theory*, 8.

published only two years after Shannon and Weaver's book.[23] In this paper, the author argued that the process of scientific measurement (for instance, measuring the temperature of a substance) could be understood by Shannon's theory as a process of communication between the observed system as information 'source' (i.e. the temperature of a substance) and the observing agent as 'destination.' In this analogy, (i) the 'message' would correspond to the value of the measured property (e.g. 147º Kelvin degrees), (ii) the 'transmitter' would correspond to the measuring device (e.g. a thermometer), and (ii) the 'receiver' would correspond to the indicator of the measured value (e.g. a scale of values in the thermometer), so that the measuring procedure would act as a communication channel between the observed system and the observing agent. Interestingly, such analogy is conceptually based on assuming the identification (previously suggested by Weaver and von Neumann) between information Shannon H and statistical-mechanical entropy S. Rothstein's analogy[24] showed how to understand this identification: the greater the capacity of the observer to distinguish different microstructures of a system in a measure, the lower the entropy and therefore the lower the information content of that measure. In short, Rothstein argued that Shannon's communication theory would find a direct application in the field of mediated interaction between observers and observed systems, shedding light on the functioning of scientific measurements.

Figure 1. Rothstein's Analogy between Communication and Measuring[25]



[23] See Rothstein, "Information, Measurement," 171-175.

[24] Rothstein's proposal would later influence more sophisticated theoretical informational proposals to scientific measurement such as that of L. Finkelstein ("Representation by symbol systems as an extension of the concept of measurement," *Kybernetes* 4, 4 (1975): 215–223), who recognized that Shannon's formalism only provided the syntactic architecture of the proposal, while measurement theory would add the strictly semantic component. However, many of the deficiencies of his proposal are already found in Rothstein's original paper (Rothstein, "Information, Measurement").

[25] Rothstein, "Information, Measurement," 171-175.

As far as the objective of this paper is concerned, this Rothsteinian naturalist proposal implicitly attempted against the non-semantic dogma that Shannon defended with respect to his proposal. If we analyze in detail this communicative-theoretical model of scientific measurement, the amount of information Shannon transmitted by the values of the measured properties ends up possessing certain semantic properties alien to the original formal apparatus. Firstly, this semantic character implicitly added to the concept of Shannon is reflected in the fact that observer A acquires communicatively or medially information about (i.e., with meaning or referring to) the microstructure of the observed system B. Secondly, scientific measurement practices as communication events end up somehow locating Shannon information in certain mental states of the observer at the end of the day. That is, by reducing all the macroscopically indistinguishable microstructures of a system through particular measurement,[26] not only does the amount of Shannon information compatible with the measurement result decrease, but also the information (in its ordinary, semantic-epistemic sense) that the agent possesses about the microstructure of the observed system. But as Mari points out, this analogy between communication and measurement becomes inconsistent because every message can be known independently of its transmission, while the state of the measured system could not be known independently of its measurement.[27] Therefore, in seeking to apply the Shannon concept of information to scientific practice, Rothstein would imply an implicit and conceptually inconsistent conflation of the theoretical-communicative (non-semantic) technical sense with the usual content of the term 'information' (semantic-epistemic).

Another of the authors who followed and even brought to its final theoretical consequences this intellectual path traced by Rothstein was the French-American physicist Léon Brillouin, who developed a systematic theoretical-informational reformulation of statistical mechanics based on the identification between the

[26] As J. Wicken ("Entropy and information: Suggestions for common language," *Philosophy of Science* 54, 2 (1987): 176-193) points out, this way of applying the concept of Shannon information in the field of statistical gas mechanics erroneously assumes that the concept of 'state' in statistical mechanics (distinguishable between 'micro-state,' a set of microscopic-molecular values that determine a system, and 'macro-state,' a set of macroscopic and measurable values on a system) could also be applied in the context of communication theory. This is not possible, precisely because while physical states are sensitive to temperature changes in the environment, theoretical-communicative states are not.

[27] L. Mari, "Notes towards a qualitative analysis of information in measurement results," *Measurement* 25, 3 (1999): 183–192.

technical notion of information à la Shannon and the negative quantities of entropy[28] or 'negentropy:'

> A more precise statement is that entropy measures the lack of information about the actual structure of the system. (...) Since any of these [indistinguishable] different microstructures can actually be realized at any given time, the lack of information corresponds to actual disorder in the hidden degrees of freedom (...) The connection between entropy and information was rediscovered by Shannon, but he defined entropy with a sign just opposite to that of the standard thermodynamic definition. Hence what Shannon calls entropy of information actually represents negentropy.[29]

To defend that the informational measure of Shannon entropy would have a real use in the field of statistical mechanics, this measure should necessarily appeal to the ability of the observer to distinguish observationally between different microscopic structures of physical systems. In this way, Shannon entropy would not simply measure the information encoded in the micro-states of the observed system but would properly measure the lack of semantic-epistemic information that an observer possesses "about the actual structure of the system." At this point we defend, following Earman and Norton,[30] that Brillouin implicitly employed the technical notion of Shannon entropy in the physical context of statistical mechanics as the usual (and therefore semantic) meaning of the notion 'information.'

As we have just seen in Rothstein's and Brillouin's cases, this tendency to implicitly semantize (note that none of them explicitly defends an alternative semantic-sensitive concept) Shannon's non-semantic informational measure in order to apply it in scientific contexts was particularly remarkable during the first years of popularization of communication theory. One of the first authors to become aware of this phenomenon was the philosophers Rudolph Carnap and Yehoshua Bar-Hillel, who were developing during that time a strictly semantic information theory based on their program of inductive logic. In the presentation of their theory at the famous Macy Conferences (series of lectures organized since the early 1950s about areas surrounding cybernetics and information theory) around 1951, these authors pointed out how many authors of the time exploited the fashionable Shannon entropy to their advantage as if it were a function sensitive to meaning or semantic content:

---

[28] Brillouin, *Science and Information*.

[29] Brillouin, *Science and Information*, 160-161.

[30] John Earman and John Norton, "Exorcist XIV: the wrath of Maxwell's Demon. Part II. From Szilard to Landauer and beyond," *Studies in History and Philosophy of Modern Physics* 30, 1 (1999): 1–40.

> It has, however, often been noticed that this [semantic] asceticism is not always adhered to in practice and that sometimes semantically important conclusions are drawn from officially semantics-free assumptions. In addition, it seems that at least some of the proponents of communication theory have tried to establish (or to reestablish) the semantic connections which have been deliberately dissevered by others.[31]

Interestingly, the very name 'information theory' was one of the main sources of confusion when it came to attributing (unintentionally and implicitly) the strong semantic character of the everyday notion of 'information' to Shannon's technical concept. As Kline shows,[32] the adoption of the name 'information theory' to refer to Shannon's communication theory occurred progressively during the first years of popularization due to the previous existence of information theories in the British sphere, such as Fisher's or Gabor's. The confusion generated by that name could be illustrated in Shannon's following comment to one of the attendees of the Eighth Cybernetics Conference in 1951, the anthropologist Margared Mead, who criticized the use of a technical notion of information that was far from its ordinary meaning: "I wanted to call the whole of what they called information theory signal theory," he said later, "because information was not yet there. There were 'beep beeps' but that was all, no information."[33]

It was precisely Shannon himself who most actively defended during the explosion of popularity of his proposal that (1) his theory was fundamentally a formal-syntactic theory about the statistical analysis of signal transmission, so that its extension to other scientific domains would not be at all immediate; and (2) that his particular concept of information (measure of entropy) was a mathematical function whose values were independent of the possible semantic values of the elements on which it is defined. This will not prevent the scientific community during the 1950s from continuing to misapply Shannon's fashionable theoretical apparatus and misinterpret his concept of information to address disciplinary problems that were not at all related to the transmission of signals through communication channels, as we have pointed out in the case of Rothstein and Brillouin. This led to what Shannon himself called the 'scientific bandwagon' after his theoretical proposal, which was set in motion not because of the results obtained in its multiple applications[34] but in order to take advantage of its growing popularity:

---

[31] Carnap and Bar-Hillel, "An Outline."

[32] Kline, *The Cybernetics Moment*, Chapter 3.

[33] This quotation of Shannon can be found in J. Gleick, *The Information: A History, a Theory, a Flood*. (New York: Pantheon Books, 2011).

[34] "The information theory approach to SM has not, to my knowledge, led to any concrete results" (Amnon Katz, *Principles of Statistical Mechanics: The Information Theory Approach* (San

> Information theory has, in the last few years, become something of a scientific bandwagon (…) Our fellow scientist in many different fields, attracted by the fanfare and by the new avenues opened to scientific analysis are using these ideas in their own problems (…) It will be all too easy for our somewhat artificial prosperity to collapse overnight when it is realized that the use of a few exciting words like *information*, *entropy*, *redundancy*, do not solve all our problems.[35]

4. Toward a Technical Concept of Semantic Information to apply in Science

We have just seen how there was an important trend within the information bandwagon in the 1950s to implicitly semantize the Shannon information measure. On the other hand, certain authors (aware of the impossibility of using this statistical concept as a measure of meaning) had the intellectual pretension of developing a theory of semantic information, where a measure was postulated that was sensitive to the meaning of the structures considered. Before evaluating the greater or lesser success of its application in specific scientific fields (paradigmatically, the measurement-observation[36] in statistical mechanics), let us first proceed to explore the main proposals of semantic information.

4.1. Bar-Hillel & Carnap: Statistical Semantic Information

As mentioned in the previous section, Carnap and his collaborator Bar-Hillel developed during the 1950s the first quantitative-statistical theory of semantic information.[37] As mentioned by the authors, this theory was framed within the broad Carnapian project to lay the conceptual foundations of inductive logic as a central methodology in the development of the empirical sciences. Unlike the theoretical proposal of Shannon used by Rothstein and Brillouin to describe the measurement processes in statistical mechanics, the theoretical proposal of Carnap and Bar-Hillel was in fact designed to illuminate how our knowledge works in scientific practices such as measurement. Broadly speaking, one of the main objectives of Carnap and Bar-Hillel's theory of semantic information is to statistically measure how the amount of 'semantic information' (understood as information about a physical property) encoded in the output of an experimental measurement E could alter our initial knowledge by decreasing the space of possible hypotheses H compatible with that scenario. Note that this space of possible hypothesis H would

---

Francisco: Freeman, 1967), i).

[35] Claude Shannon, *The Bandwagon* (1956), in *Collected Papers*.

[36] In this paper we assume that all scientific measurement is an observational process, but not the other way around.

[37] See Carnap and Bar-Hillel, "An Outline."

be partitioned by a conceptual framework L into different particular hypotheses. For example, the L_{SM} conceptual framework of statistical mechanics partitions the space of possible (non-measurable-observable) hypothetical H microscopic configurations of a substance with respect to the (measurable-observable) temperature values of that substance, so that obtaining 189º Kelvin in a thermometric measurement would reduce the space of possible microscopic configurations to a particular subset of this space. From these theoretical elements it could be formulated its measure of semantic information (represented in the function 'inf') contained in a proposition *i*:

$$\inf(i) = -\log m(i) \qquad (2)$$

With this quantitative-statistical measure of the amount of semantic information inf(*i*) proposed by Carnap-Bar-Hillel (defined as the negative logarithm of the number of events m referred by the content of proposition *i*) it would be possible to evaluate the particular conceptual framework L that would maximize the information provided by an experimental data in the selection of the most appropriate hypothesis. For example, from the conceptual framework of L_{SM} statistical mechanics, the same proposition *i* about the temperature of a substance would provide more information than from the framework of L_{TD} thermodynamics: namely $inf_{SM}(i) > inf_{TD}(i)$. This is precisely because the number of events (hypothetical microconfigurations of the substance) to which the 189º Kelvin refers in a thermometric measurement are reduced more from the framework of statistical mechanics than from the framework of thermodynamics.

In this line, the authors[38] clearly differentiated between the amount of information 'inf(i)' of a logically structured statement or proposition i, which depended on the negative logarithm of the possible alternatives (as with Shannon entropy); and the content measure 'cont(i)' of a proposition i, which refers to the number of compatible events indicated by i. Thus, while in the domain of the function cont(i) (where i = 'the measured temperature of this substance is 189º degrees Kelvin') the set of microscopic events referred to by the value of 189º K included in proposition i would be included, the domain of the function 'inf(i)' would properly cover the number or quantity of events referred to by this same proposition. Contrary to Shannon's theoretical proposal, Carnap and Bar-Hillel argued that their theory was constitutively based on the evaluation of the semantic content of the vehicles of information as symbols-messages, or properly in their case, declarative sentences or even logical propositions:

> Prevailing theory of communication (or transmission of information) deliberately

[38] See Carnap and Bar-Hillel, "An Outline."

> neglects the semantic aspects of communication, i.e., the meaning of the messages. This theoretical restraint is, however, not always adhered to in practice, and this results in many misapplications. The theory outlined here is fully and openly of a semantic character.[39]

4.2. MacKay: Descriptive Information

When Shannon presented his communication theory at the first London Symposium on Information Theory in 1950, he realized that a significant number of UK scientists (what some historians such as Kline[40] called the English School of Information Theory) used the term 'information theory' to refer to a disciplinary field that transcended that of Shannon's theory: "They adopted the name Information Theory to refer to a broader concept of information than that held by Shannon."[41] If we have to point out the differential feature of the English theories of information as opposed to Shannon's North American alternative centered on the transmission of symbols, this would undoubtedly be the creation of a concept of information focused on the clarification of scientific practices: "the concept of Information has wider technical applications that in the field of communication engineering. Science in general is a system of collecting and connecting information about nature."[42] In this direction, the physicist Donald MacKay sought to develop (during a period spanning from the early 1950s to 1969, the date of publication of his seminal work) a general theory of information that would account for the system of collecting and connecting information that was scientific practice.[43]

From his early interest in the problem of the objective limits of scientific measurement "the art of physical measurement seemed to be ultimately a matter of compromise, of choosing between reciprocally unrelated uncertainty,"[44] the young scientist MacKay will intellectually evolve towards the question of how we should quantify (he will propose the term 'information quantum,' analogous to the 'atomic propositions' of Wittgenstein's Tractatus[45]) and understand information as it naturally appears in the context of measurement in actual scientific practices. In this way, MacKay's information theory will not be detached from particular physical

[39] See Carnap and Bar-Hillel, "An Outline," 1.
[40] See Kline, *The Cybernetics Moment*, 104-111.
[41] Kline, *The Cybernetics Moment*, 105.
[42] Kline, *The Cybernetics Moment*, 206.
[43] See MacKay, *Information, Mechanism.*
[44] See MacKay, *Information, Mechanism*, 1.
[45] "Although (…) Wittgenstein himself had now repudiated the atomistic approach of the Tractatus, the field of scientific measurement seemed well suited to logical treatment on these" (*Ibid*, 2)

theories, as in the case of Shannon's proposal, but it will otherwise emerge naturally (or be already incorporated) from the physical theory itself employed by the agent to obtain knowledge via measurement processes.

Interestingly, MacKay incorporated into his ambitious information theory (i) the Fisher measure of statistical information or 'metric information' (renaming it 'metron') defined as the reciprocity of a parameter O about a variable X; and (ii) the Gabor measure of physical information based on minimum units of phase-frequency-time volume or 'structural information,' renaming the latter as 'logon.' Both measures of information make up what this author calls 'descriptive information,' due to their role in being used by scientific observers to describe the phenomena observed through certain measurement processes. For example, the possible values of temperature-parameter O that we can measure on a substance provide the agent-observer with certain metric and semantic information about (i.e., descriptively or intentionally) the hypothetical variables of position and molecular velocity that make up the microstructure of a substance. In the same way, all the vectors of position and velocity that determine the exact 'micro-state' of a gas at a particular moment provide the agent-observer with certain structural information that accurately describes the microscopic state of the system.

In addition to the metric (Fisher's) and structural (Gabor's) information, MacKay[46] includes in his ambitious theory the Shannon entropy, redefining it as "selective information," precisely because its value depends on how the units-symbols with which the messages to be transmitted are "selected." Recognizing this informational measure as merely syntactic, this author will also argue in his *Information, Mechanism, and Meaning* that the concepts of metric information and structural information (both grouped under the name of descriptive information) possess a semantic character derived from the intentionality of these agent-observers: “It appears from our investigations that the theory of information has a natural, precise, and objectively definable place for the concept of meaning.”[47] However, MacKay seems to suggest (although it is not very clear) that the semantic character of these concepts and measures of information would depend mainly on the capacity of the agent-observers to employ them as representational tools of the objective phenomena.

[46] See MacKay, *Information, Mechanism*.

[47] See MacKay, *Information, Mechanism*, 93.

4.3. Dretske: Semantically-Enhanced Shannon Information

So far, we have explored (i) a concept of semantic information developed in parallel with Shannon's measurement (i.e., Carnap-Bar-Hillel information) and (ii) another concept developed as a semantic alternative to this same information measure during the 1950s and 1960s, i.e., MacKay's descriptive metric and structural information. I will now detail a third strategy to obtain a technically defined notion of semantic information, based mainly on the theoretical-formal supplementation of Shannon's measure in such a way that it is effectively capable of capturing certain types of semantic relations existing in the messages transmitted communicatively. Among all the concepts that we can find within this group, one of the most important and influential within literature is that developed by the philosopher Fred Dretske in the early eighties to develop a naturalistic theory of human perception based closely on Shannon's theory of communication.[48]

Although this author recognizes from the beginning the Shannonian dictum that its entropy measure does not capture the plausible meaning of a sequence of symbols, it is also part of Weaver's suggestion that certain engineering considerations might be relevant to measuring certain semantic relationships (see Section 2). To this end, communication theory should be technically and theoretically supplemented to derive semantic relations from the statistical correlations between the occurrence of symbols and the occurrence of events: "the underlying structure of this [communication] theory, when suitably supplemented, can be adapted to formulate a genuinely semantic theory of information."[49] As we pointed out at the beginning, one of the main reasons why Shannon's communication theory is insensitive to the semantic content of messages is precisely because its measurement entropies not a measure of the amount of information of a particular sequence of signs (as one would expect from a measure of semantic information) but a statistical assembly of possible messages. Dretske proposes to develop a measure I of the information transmitted by an individual event y_i (although the author prefers to speak of 'signals' and 'states of affairs' to refer to the events they refer to and to the events referred to, respectively) to solve this theoretical problem, representing the amount of information that a sequence of particular events y_j transmits about another event or particular state of affairs x_i, thus assimilating the intentional character or 'aboutness' that must characterize any robust semantic information concept. This measure could be formulated by the following mathematical expression:

[48] See Dretske, *Knowledge and the Flow.*

[49] See Dretske, *Knowledge and the Flow*, x.

$$I_{xi}(y_j) = -\log p(x_i) - H\, p(x_i|y_j) \quad (3)$$

where $x_i \in \{x_i\}$, i = 1, . . ., m; $y_j \in \{y_j\}$, j = 1 ..., n. This measure would represent[50] the amount of information generated by the existence of certain signal-events x_i about the occurrence of other events y_j, which is determined by the logarithm of the improbability of occurrence of the signal-event x_i minus the H-encoded unexpectedness of the conditional probability distribution between the event referring to x_i and the referred event y_i. Therefore, its measure of semantic information (3) would definitely depend on the non-semantic information measure of Shannon entropy (1), deriving semantic relations from the statistical correlations between the two events. With this measure of the amount of information contained in a single event, Dretske would have the necessary resources to define the semantic information that would be contained in a particular signal S about or about a certain event q:

$$\text{A signal S contains the information that q def} = p\,(q|S) = 1 \quad (4)$$

That is, an event that acts as an S signal (that is, the referenced event) would contain information about another q event (or referenced event) if there is a complete correlation between the S signal and the referenced event. This would capture the modal-counterfactual amount of semantic information; namely, that each time the S signal occurs, it necessarily depends on the fact that the p event to which it refers has also occurred. Dretske explains this modal feature by underlying that the event referred q acts properly as the cause of the signal S, wherein S in turn constitutes the effect of that causal connection. Thus, the fact that a signal S has semantic information about q would be based on the existence of an asymmetric causal relationship between q and S, which for reasons of extension we will not go into further. The key here is that, according to the Dretskean proposal, the semantic information would constitute a statistical property derived from the statistical correlation between events, where its quantity is statistically determined by the conditional probability of (4): while the value 1 would represent that the signal S has the maximum semantic information of which q, the value 0 would represent that the signal S has no information about q.

[50] See Dretske, *Knowledge and the Flow*, Section 2.3 and Christopher Timpson, *Quantum Information Theory and the Foundations of Quantum Mechanics* (Oxford: Oxford University Press, 2013).

5. Against Semantifying Quantitative and Shannon Information

We have just described the main proposals for the development of semantic information concepts in relation to Shannon's non-semantic notion. First, Bar-Hillel and Carnap postulated a measure of semantic information formally identical to Shannon entropy, defined as the logarithm of the improbability of a proposition i. Second, MacKay proposed to reinterpret the representational tools we find in scientific practices as types of 'descriptive information' about the referred phenomenon. Finally, Dretske developed a measure of semantic information based on the statistical correlations between S-signals and state of affairs q that we can extract from Shannon's formalism. Next, our task will be to assess whether these proposals can help to illuminate the knowledge of agents in scientific practices in a more satisfactory way (in terms of conceptual consistency and interpretative coherence) than with Shannon's proposal, as we saw in the case of Rothstein and Brillouin.

5.1. Meaninglessness of Carnap-Bar-Hillel Semantic Information in Real Scientific Practices

The semantic information theory of Carnap and Bar-Hillel[51] was directly conceived to analyze the logical architecture of scientific practices, so in principle we could assume that their concept of information can satisfy this role. Now I will argue that this is not the case, since this proposal presents an important conceptual problem that makes it unsuitable for this task. Namely, as in the case of Shannon's proposal, his measurement of the information content of a proposition i refers to those events that are compatible with that content. That is, a proposition will have the highest degree of semantic-informational content when any possible event will be compatible with (or satisfy) said proposition i, thus we will find what is known as a tautological or trivially true proposition. In the same way, a proposition i will be minimally informational when no event is compatible with its content, thus finding ourselves with a contradictory proposition. However, the fact that the maximum value of the semantic information function cont(i) is found in a tautological proposition could indicate that this proposition is the one that provides more information (in the sense of 'knowledge') about the domain of events considered. But what happens is precisely the opposite. While according to Carnap and Bar-Hillel's proposal, tautological propositions are those with the maximum informational content, they are also those that inform us the least (precisely because of their triviality) about the domain in question. And in the same way with

[51] See Carnap and Bar-Hillel, "An Outline."

contradictory propositions: although they have the minimum informational content, they are the ones that inform us the most (perhaps in excess) about the domain of events. This is precisely what Floridi calls the 'Carnap-Bar-Hillel paradox' of semantic information,[52] which could be synthesized in the idea that the proposition that has the most information about a domain is the one that informs us the least about this same domain.

As might be expected, this has terrible consequences when it comes to analyzing scientific practices. Let us suppose that we consider the proposition i_1 = 'the molecules of this gas possess a certain position and velocity.' According to the proposal of Carnap and Bar-Hillel, the proposition i_1 would have the maximum informational semantic content with respect to the domain of microscopic configurations of the gas, since virtually any possible microstructure of the gas would be (again, trivially) compatible or satisfy such a proposition. However, this same proposition i_1 would not provide any significant information to the agent about the actual microstructure of the gas, since for any sufficiently competent agent it would be completely trivial or non-informative the fact that the molecules of a gas have a certain position and speed. What would be significantly informative for the agent would be a proposition (with much less informational content than a trivial proposition such as i_1) in which the position and speed of the molecules would be determined. Note that in the same way, a contradictory proposition with the minimum informational content (according to the proposal of Carnap and Bar-Hillel) such as i_2 = 'the molecules of this gas possess and do not possess a certain position and velocity' would provide as little information to an agent as a trivial proposition i_1. In any case, the measurement of the semantic information content of Carnap and Bar-Hillel does not satisfactorily represent the degree of informativity that this semantic content would provide to an agent with respect to the domain of phenomena considered, even describing the behavior of quantities of information in a substantially paradoxical way. For these reasons, the proposal of Carnap and Bar-Hillel does not constitute either the development of a concept of semantic information robust enough[53] to significantly illuminate the dynamics of knowledge in scientific practices.

[52] See Floridi, *Philosophy of Information*, Chapter 3.

[53] Floridi (*Philosophy of Information*) described the theoretical proposal of Carnap and Bar-Hillel ("An Outline") as a weak semantic information theory, precisely because of its inability to be sensitive to epistemically relevant values within a particular context.

5.2. Meaninglessness of Dretskean Semantic Information in Real Scientific Practices

Undoubtedly, Dretske's proposal constitutes one of the most important attempts to obtain a technically sophisticated notion of semantic information. However, I will now show how this would still not be robust enough to account for the dynamics of knowledge in scientific practices. We can conceive of two types of objections to the Dretskean program, namely, technical objections and theoretical objections. As for the former, we can follow Timpson[54] in criticizing the concept of semantic information $Ix_i(y_j)$ (3) as a highly deficient measure of the information that x_i contains about y_i. The central reason is that the first term of the measure $-\log p(x_i)$ (formally identical to Carnap and Bar-Hillel's semantic information measure) would be completely independent of the second term $H\, p(x_i|y_j)$:

> For example, our uncertainty in x_i given y_j might be very large, implying that we would learn little from y_j about the value x_i, yet still the amount said to be carried by y_j about x_i, under Dretske's definition, could be arbitrarily large, if the surprise information$_t$ of x_i dominates. Or again, the channel might be so noisy that we can learn nothing at all about x_i from y_j -the two are uncorrelated, no information can be transmitted- yet still $Ix_i(y_j)$ could be strictly positive and very large (if the probability of x_i is sufficiently small). This is sufficient to show that $Ix_i(y_j)$ is unacceptable as a measure. The hoped-for link to information theory is snapped.[55]

That is, let us assume in a context of scientific practice that events y_i represent the possible microscopic (non-observable) configurations of a gas, as well as events y_i represent the possible observable temperature values (e.g., marks on a thermometer) that we can extract by thermometric measurement from a gas. In this scientific measurement scenario, an agent could possess a high Shannon-encoded uncertainty $H\, p(x_i|y_j)$ on the actual microstructure of the gas y_i given a set of microconfigurations compatible with a particular temperature value x_i, which represents that the agent cannot recognize which is the actual microstructure only from her knowledge of its temperature value (i.e. 189º Kelvin). At the same time, according to Dretske,[56] the amount of information that the 189º Kelvin of the gas contains about the actual microstructure of the gas can be arbitrarily high, and therefore it would be independent of the uncertainty that the agent possesses about such microstructure.[57] Consequently, the amount of semantic information $Ix_i(y_j)$

[54] Timpson, *Quantum Information Theory*, 40.

[55] Timpson, *Quantum Information Theory*, 40.

[56] See Dretske, *Knowledge and the Flow*, x.

[57] "Unfortunately, the quantity $Ix_i(y_j)$ cannot play the role of a measure of the amount of information that y_j carries about x_i. To see this we need merely note that the surprise [Shannon] information$_t$ associated with x_i is largely independent of the uncertainty in the conditional

could not satisfactorily capture the amount of information that the thermometric values of the gas carry on the actual microstructure of that same gas, and with even less technical consistency would represent the amount of informativeness of the thermometric signal for the competent agent with respect to the molecular properties of this substance.

As for the theoretical deficiencies that we could object to Dretske's proposal of 'information that', we will find its excessive idealization of the informational dynamics. In the first place, in order to specify how much information content would be linked to the occurrence of the event we should determine the domain of different possible events. Illustratively, to know how much information the 189º K event provides us about the actual microstructure of the gas we should know in advance how many different possible microstructures are compatible with that thermometric signal. However, it would be reasonable to think that it is not possible to specify the domain of different events in realistic contexts of application of the Dretskean apparatus. In fact, if we consider the measurement practices of classical statistical mechanics, the fact that the number of distinct micro-stages of a gas is uncountably infinite (precisely because the classical phase space of a gas is continuous) makes it practically and conceptually impossible to determine the domain of events in this particular case. Therefore, Dretske's measure will not only be technically deficient because of its combination of two independent terms in (3), but also theoretically ideal, since it could only be applied in extremely simple and uninteresting scenarios with respect to our understanding of sufficiently realistic scientific practices.

6. Defending a Neo-MacKayian Meaningful Use of Information Concepts in Science

Unlike the proposals of Carnap-Bar-Hillel and Dretske that we have previously analyzed, MacKay's theory of semantic-descriptive information does not depend on the development of a measure of its own, but on a kind of interpretative exercise carried out on existing representational structures (i.e. reciprocal quantities of variance, volume of phase space) within real scientific practices.[58] This fact may be a vice or a virtue: a vice if the interpretative exercise is not sufficiently robust or well-defined theoretically (as is unfortunately the case with MacKay), and a notable virtue because of the enormous descriptive potential that a well-defined proposal developed in this direction would possess. In this section I will attempt to carry out this task.

probability distribution for x_i given y_j" (Timpson, *Quantum Information Theory*).

[58] MacKay, *Information, Mechanism*.

The ambitious information theory that MacKay originally presented at the First London Symposium on Information Theory was severely criticized by his contemporaries during the 1950s, right in the middle of the maelstrom of Shannon's bandwagon. As Kline reminds us, Colin Cherry (also from the English School of Information Theory) argued in 1956 that Bar-Hillel and Carnap's semantic information theory was the "only investigation of which your author is aware, into the possibility of actually applying a measure to semantic information," an implicit critique of MacKay.[59] This disregard for MacKay's information theory, with no repercussions in the literature, could be justified on a theoretical level by the enormous indefiniteness of his proposal and by the vagueness of his definitions, as I will proceed to show. As far as we are concerned in this paper (i.e., the application of informational concepts in mediated practices in science), the author states that "in physics, of course, the notion of descriptive information-content is often more useful than that of selective [Shannon] information-content. A physicist may want to know how much theoretically available information is being wasted in a given microscope."[60] As this quotation shows, MacKay defends that his concept of descriptive information (Section 4.2) is a truly significant alternative to the informational concept of Shannon entropy (on which the proposals of Carnap-Bar-Hillel and Dretske, respectively, rest indirectly or directly) to be applied in an epistemically useful way in the context of scientific practices.

But what does the author mean by 'theoretically available information' about a physical system? If we analyze the text in depth, we can arrive (seventy-seven pages later) at something similar to a definition: "A particular message may now be pictured as selecting a particular region, which may be identified by the vector (or a distribution of vectors) linking it to the origin. The meaning of the message is then represented by the orientation of this vector, relative to the vector basis."[61] Therefore, from this proposal the semantic information that the agent-observer has on the observed system is a resource that is defined by the theory used, which determines what MacKay calls the 'representational space' built by the vectors that parameterize the system under consideration. However, here we argue that it is not necessary to introduce an additional space to that already used in real scientific practice. In the case of statistical mechanics that we have used throughout this paper, the phase space of the system itself (abstract space that encodes all the possible values of position and molecular speed of the system) would play the role of

[59] See Kline, *The Cybernetics Moment*, 111.
[60] See MacKay, *Information, Mechanism*, 15.
[61] See MacKay, *Information, Mechanism*, 92.

representational parameterization of statistical-mechanical information about the system.

The main flaw of MacKay's proposal lies precisely in seeking to quantify the meaning of a description, represented in a vector within the phase representational space, by means of the orientation of this vector. That a vector-description has a certain orientation is a conventional-arbitrary fact (subject to how we design the representational space) and in no way exhausts the semantic information encoded in such a description within the system's representational space. If this were the case, then after performing a thermometric measurement we would obtain a new statistical-mechanical description whose meaning would be the opposite of the description prior to the thermometric measurement or observation, precisely because the rules of vector calculus state that a difference in vectors changes the orientation of the resulting vector (see Figure 2).

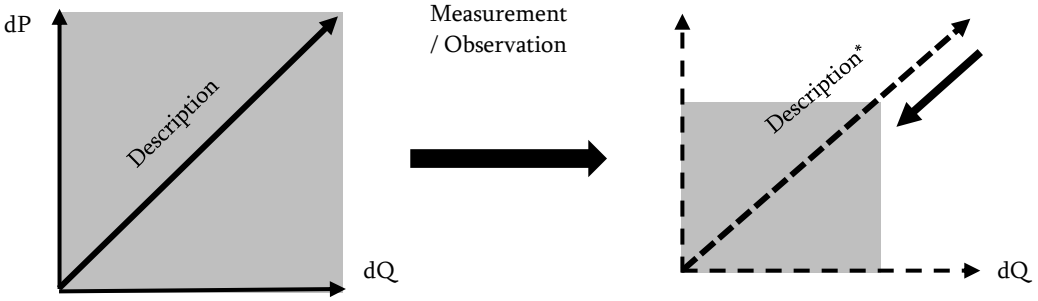


Figure 2. MacKay semantic information defined in representational space. The phase space represents information about the molecular position (P) and velocity (Q), where the orientation of the vector-description represents its meaning. Therefore, according to the rules of vector calculation, the meaning of the description* (or its semantic information) after performing a measurement on the system would be the opposite of meaning of the previous description.

However, it would be absurd to think that the semantics of this sort of descriptions made in a realistic measurement scenario closely depends on the rules of vector calculus: these rules fix how we can construct vector representations of a phenomenon, but they do not properly fix the meaning of those theoretical elements we use to describe them. Contrary to MacKay's concrete proposal but following his framework, we propose to understand the semantic information of a statistical-mechanical description as the way in which theory delimits the way in which representational resources (i.e. micro-states as 'phase points', macro-states as 'phase

regions') represent target phenomena. Illustratively, statistical mechanics[62] (not vector calculus or any other mathematical theory employed) theoretically determine (i) that each point of the representational phase space encodes exact semantic information about the position and velocity of each and every molecule that composes a particular substance; and (ii) that each region of the representational phase space would encode semantic information about the macroscopic properties (temperature, pressure, and so on) on this substance.

It is the theoretical apparatus of statistical mechanics that specifies how these two types of semantic information (non-observable information about molecular components, observable information about substances) are interconnected in a context of scientific practice such as measurement; and in no case could this be specified in a realistic formal-theoretical apparatus of the statistical theory of signal transmission (against Rothstein[63] or Dretske[64]). For example, in Boltzmann's statistical-mechanical formalism, each piece of observable semantic information (e.g., the 189º K value) would be theoretically associated with a counterfactual set of non-observable semantic information (e.g., semantic information about two observationally indistinguishable microscopic configurations). In this way, an agent competent with the practices of statistical mechanics could extract significant semantic information about the real micro-state of the system from the macroscopic values of observable properties (e.g. temperature) measured, the latter associated with sets of micro-state counterfactually compatible with that macroscopic value.

Another of the main advantages of our Neo-MacKayian proposal is that it would allow us to incorporate the theory of scientific measurement recently defended by van Fraassen, where the qualitative concept of semantic-epistemic information would play an essential role. For van Fraassen[65], the possible measurable states of a physical system are represented in a parameterized space. Thus, a particular measurement would inform us about the state of the target system by reducing the parameterized space of possible values to a subsystem. In our case of study of the measurements in statistical mechanics, the parameterized space is none other than the phase space of possible microscopic configurations of the system, which would be reduced to a particular sub-region (sub-set of micro-states) after a macroscopic measurement on the system (see Figure 2). In this way, the meaning of

[62] Note that to defend this particular thesis we do not need to support any particular conception of scientific theories.

[63] Rothstein, "Information, Measurement."

[64] See Dretske, *Knowledge and the Flow*, x.

[65] Bas van Fraassen, *Scientific Representation: Paradoxes of Perspective*, (Oxford: Oxford University Press, 2008), 141-185.

the semantic information about a gas contained in the phase space would not change radically (as suggested by MacKay[66]) after a thermometric measurement on this gas, but would increase its epistemic relevance for the agent-observer with respect to the recognition of the microscopic configuration of the measured system.

Unlike the quantitative-semantic information proposals of Carnap-Bar-Hillel and Dretske, MacKay's information-theoretical framework (although not so much his particular proposal) would allow us to develop a robust naturalistic theory about how scientific agents obtain semantic information without the need to postulate forced analogies with signal transmission processes or without relying on an information concept formally close to Shannon's. Our proposal has been an apology for the use of the qualitative concept of 'semantic information' within a robust philosophical analysis of scientific practices (in our case we have focused on measurement) against the hegemonic tendency since the mid-twentieth century to quantify-model the semantic information of scientific agents with the formal basis of Shannon's communication theory.

7. Conclusion

In this paper we have shown how the first attempts to apply the concept of Shannon information in the field of scientific practice (e.g., Rothstein's and Brillouin's[67]) failed precisely because of the lack of semantic character of this notion. Therefore, we have explored the main proposals of semantic information theories developed on Shannon's quantitative notion applied in the context of science (e.g. Carnap and Bar-Hillel's, and Dretske's[68]), arguing later that each of them lacks the theoretical capacity to satisfactorily account for the information dynamics between observers and systems observed in realistic scientific contexts. Finally, we have defended not the particular proposal but the informational framework of MacKay,[69] based on the informational interpretation of the representational resources of real scientific practices and arguing that it would be robust enough to be satisfactorily employed in philosophical analyses of this field. The following quote from Colin Cherry clearly shows the spirit of what we have sought to defend throughout this paper:

> Questions of extracting information from Nature and of using this information our models or representations lie outside communication theory - for an observer looking down a microscope, or reading instruments, is not to be equated with a

[66] MacKay, *Information, Mechanism*, 92.

[67] See Rothstein, "Information, Measurement, and Quantum Mechanics," *Science* 114 (1951): 171-175 and Brillouin, *Science and Information*.

[68] Dretske, *Knowledge and the Flow,* x.

[69] MacKay, *Information, Mechanism.*

listener on a telephone receiving spoken language. Mother Nature does not communicate to us with signs or language. A communication channel should be distinguished from a channel of observation.[70]

[70] E.C. Cherry "On Validity of Applying Communication Theory to Experimental Psychology," *British Journal of Psychology* 48 (1957): 176-188, quotation in Kline, *The Cybernetics Moment*, 111.

EPISTEMIC NORMS, THE FALSE BELIEF REQUIREMENT, AND LOVE[1]

J. Spencer ATKINS

ABSTRACT: Many authors have argued that epistemic rationality sometimes comes into conflict with our relationships. Although Sarah Stroud and Simon Keller argue that friendships sometimes require bad epistemic agency, their proposals do not go far enough. I argue here for a more radical claim—romantic love sometimes requires we form beliefs that are false. Lovers stand in a special position with one another; they owe things to one another that they do not owe to others. Such demands hold for beliefs as well. Two facets of love ground what I call the *false belief requirement*, or the demand to form false beliefs when it is for the good of the beloved: the demand to love for the right reasons and the demand to refrain from doxastic wronging. Since truth is indispensable to epistemic rationality, the requirement to believe falsely, consequently, undermines truth norms. I demonstrate that, when the false belief requirement obtains, there is an irreconcilable conflict between love and truth norms of epistemic rationality: we must forsake one, at least at the time, for the other.

KEYWORDS: doxastic wronging, love, false belief requirement, epistemic partiality

1. Introduction

The epistemology of romantic love has been a largely neglected topic in the literature on epistemology and relationships. This literature has focused primarily on the epistemic demands of friendship (and the conflicts friendship poses to epistemology or deflating the conflict).[2] Love, however, poses interesting epistemic demands excluded from friendship, and these demands show deeper conflict than friendship. Loving relationships make special demands on us; we owe things to our

[1] Thanks to Mattias Iser and Hilde Lindemann for helpful comments and suggestions on previous drafts.

[2] See Sarah Stroud, "Epistemic Partiality in Friendship" *Ethics* 116, 3 (2006): 498-524; Jason Kawall, "Friendship and Epistemic Norms," *Philosophical Studies* 165 (2013): 349-370; Allan Hazlett, *A Luxury of the Understanding* (Oxford: Oxford University Press, 2013); Katherine Hawley, "Partiality and Prejudice in Trusting," *Synthese* 191 (2014): 2029-2045; Christian Piller, "Evidentialism, Transparency, and Commitments," *Philosophical Issues* 20 (2016): 332-350; Lindsay Crawford, "Believing the Best: On Doxastic Partiality in Friendship," *Synthese* 196 (2019): 1557-1593; and Sanford Goldberg, "Against Epistemic Partiality in Friendship," *Philosophical Studies* 176 (2019): 2221-2242.

lovers that we do not owe to others. Some of these are demands are epistemic. It matters, for example, what we believe about our beloved, even when it is inconsistent with the evidence or even false. The demands of romantic love require that we sometimes become bad epistemic agents, or at least I will argue.

I argue here that Sarah Stroud's account of epistemic partiality is not radical enough. Stroud offers a constraint on epistemology: if epistemic rationality precludes elements of the good life, then we have reason to rework our definition of epistemic rationality in order to include elements of the good life. I argue, however, that love, an element of the good life, sometimes requires false belief and that a plausible account of rationality cannot ignore the need for true belief. Within the context of love, we may be required to hold false beliefs about our lover—this is the *false belief requirement*. I, therefore, propose that there are sometimes irreconcilable conflicts between epistemic norms and love. Love, in short, can require irrationality. Two facets of love can pose the false belief requirement: loving for the right reasons and the demand not to doxastically wrong.

I first briefly survey concepts from modern analytic epistemology, noting that they are inseparable from the truth condition. Knowledge, we will see, requires truth. I then turn to Stroud's epistemic partiality in friendship. In this section, I outline her account and then turn to the implications for epistemology and friendship. Next, I examine two facets of love, each of which may require bad epistemic behavior. I then argue that romantic love sometimes requires that we hold false beliefs. I first turn to Neil Delaney's account of loving for the right reasons. I argue for the agent-oriented claim that the loving agent must adopt the reasons the beloved holds dear about herself. The next section argues that lovers refrain from doxastic wronging. Lovers make themselves especially susceptible to doxastic wronging. Since true beliefs can doxastically wrong, lovers sometimes are required to believe falsely. In the last section, I explore one implication of the demand for false belief: that romantic love sometimes requires irrationality.

2. The Unalterable Epistemic Norm of Truth

I briefly want to consider the centrality of truth to modern analytic analyses of knowledge and epistemic rationality. An epistemically responsible agent, according to these views, must form beliefs that are true.[3] I give examples and adaptations of

[3] This claim in controversial as there is a debate about whether truth is required for warranted belief. See, for example, Trenton Merricks, "Warrant Entails Truth," *Philosophy and Phenomenological Research* 55, 4 (1995): 841-855; Daniel Howard-Snyder, Frances Howard-Snyder, and Neil Feit, "Infallibilism and Gettier's Legacy," *Philosophy and Phenomenological Research* 63, 2 (2003): 304-327; and E.J. Coffman, "Warrant without Truth?," *Synthese* 162 (2008):

the truth condition in modern analyses of knowledge and then suggest that truth is indispensable to modern analytic epistemology. If the truth condition is indispensable to epistemology, then the false belief requirement will be opposed to epistemic norms, namely believing truly. As I suggest later, when love poses the false belief requirement, love will be opposed to epistemic norms.

Consider, firstly, the "traditional" account of knowledge—knowledge as justified true belief or JTB analysis. This account has three conditions: S believes p, S is justified in believing P, and p is true. The remarkably non-controversial condition here is the truth condition. Post-Gettier analyses of knowledge have, as far as I can tell, maintained the truth condition.

Evidentialism is a supplemental account of knowledge that follows the JTB analysis. According to evidentialism, a belief counts as justified if and only if the available evidence supports the belief in question. W.K. Clifford writes: "It is wrong always, everywhere, and for anyone to believe anything upon insufficient evidence."[4] Evidence here is so important because the evidence tracks truth.

Consider now some conditions on knowledge. Sensitivity and safety are also deeply related to true belief. Robert Nozick proposes the sensitivity condition: S is sensitive to the truth if and only if S would not hold a belief p if p were false.[5] Responding to the many counterexamples to the sensitivity condition, Ernest Sosa proposes safety. The safety condition says that in any possible world where S holds a belief p, p is true. That is, there is no possible world where p is false, and S believes p.[6] Put simply, "S would not believe p without it being so that p."[7] Safety and sensitivity, commonplace mechanisms in modern epistemology, are modal reconfigurations of the truth condition.

Consider the relevant alternatives condition on knowledge: S knows that P only if S can rule out relevant alternatives to P.[8] But not all alternatives are relevant—when, for example, someone, as I am trying to figure out where my wallet is, claims that aliens stole it. I need not rule this possibility out because this error

---

173-194.

[4] W.K. Clifford, "The Ethics of Belief," in *The Ethics of Belief and Other Essays,* ed. Tim Madigan (Amherst, MA: Prometheus), 77.

[5] Robert Nozick, *Philosophical Explanations* (Oxford: Oxford University Press, 1981).

[6] Ernest Sosa, "How Must Knowledge Be Modally Related to What Is Known?," *Philosophical Topics* 26 (1991): 373–384.

[7] *Ibid.*, 378.

[8] I have taken this formulation of the necessary condition from Georgi Gardiner, "Risk and Relevance: How the Relevant Alternatives Framework Models the Epistemology of Risk," forthcoming in *Synthese*. See, as well, David Lewis, "Elusive Knowledge," *Australasian Journal of Philosophy* 74 (1996): 549-567.

possibility is "remote." What makes an alternative relevant or "non-remote"? Some argue that an error possibility is non-remote only if it in fact obtains. That is, whether or not a possibility is in fact true is what makes it relevant.[9] This formulation, I think, largely tracks truth norms in epistemology—relevance is contingent upon what actually obtains.

These are but a few examples of the fundamentality of the truth condition in modern analyses of knowledge: knowledge "is a kind of relation to the truth."[10] Analytic epistemology, I think we can reasonably conclude, cannot compromise the truth—it is indispensable to epistemology.

I argue later that love can sometimes pose the *false belief requirement*. Being a good lover sometimes requires we adopt false beliefs about our beloved. If this is right, then, given the fundamentality of truth to epistemic rationality, there is sometimes an irresolvable dilemma between epistemic rationality and love. This dilemma demonstrates a deeper epistemic conflict than Sarah Stroud's epistemic partiality, to which now I turn.

3. Epistemic Partiality and Its Implication for Epistemology

Sarah Stroud and Simon Keller have identified that there are sometimes conflicts between friendship and epistemology.[11] Stroud says that friends owe one another epistemic partiality. I should not, for example, believe that my friend has done something that reflects poorly of his character; I owe it to him (in a manner that I do not owe to strangers) to interpret the evidence in the very best possible light. In short, friends owe one another differential epistemic practices. Such practices, however, do not go far enough; love requires something beyond epistemic partiality—false beliefs.

[9] See Lewis, "Elusive Knowledge."

[10] Jonathan Ichikawa and Matthias Steup, "The Analysis of Knowledge," in *The Stanford Encyclopedia of Philosophy*, ed. Edward Zalta, 2018.

[11] Stroud, "Epistemic Partiality;" and Simon Keller, "Friendship and Belief," *Philosophical Papers* 33, 3 (2004): 329–351. Others too acknowledge the tension between friendship and epistemology: William James, *The Will to Believe and Other Essays in Popular Philosophy* (New York: Dover Publications, 1896/1956): 1-31; and Scott Aikin, "Evidentialism and James' Argument from Friendship," *Southwest Philosophy Review* 24 (2008): 173-180. Some endorse a partialist position outside of the context of friendship: Jack Meiland, "What Ought We Believe? Or the Ethics of Belief Revisited," *American Philosophical Quarterly* 17 (1980): 15-24; Hazlett, *A Luxury*; and Piller, "Evidentialism, Transparency."

Stroud characterizes four epistemic demands of friendship. Following Sanford Goldberg, I use the following labels for the demands: *Serious Scrutiny, Different Conclusions, Interpretive Charity,* and *Reason*.[12]

First, *Serious Scrutiny*. We scrutinize negative claims about our friends: unsavory claims about our friends are harder to justify because we tend to be more skeptical when our friends are in question. If the evidence favors the unsavory claim, we "tend to devote more energy to minimizing the impact of unfavorable data than we otherwise would."[13]

Stroud also thinks we derive *Different Conclusions* when our friends are in question: Friends "draw different conclusions and make different inferences than they otherwise would" with non-friends.[14] Friends "are simply less likely to conclude that our friend acted disreputably, or that he is a bad person, than we would be in the case of a nonfriend."[15]

The third requirement is *Interpretive Charity*. We also interpret evidence against our friend more charitably than with non-friends—unsavory claims are just expressions of neutral, rather than malicious, character traits. Stroud thinks that partiality is "a matter of extending some interpretive charity to your friends than you naturally would to strangers."[16] My friend's loud behavior, I may conclude, is not obnoxious, but rather "refreshingly forthright."[17]

We, lastly, treat the fact someone is a friend as a *Reason* when we believe about them. Stroud writes: "The good friend's reason for adopting these differential epistemic practices seems to be simply that the person in question is her friend. But that someone is your friend is not a relevant epistemic reason…to form different beliefs about him than you would about anyone else."[18] These are the epistemic demands of friendship; Stroud thinks one must follow these demands to be a good friend.

Stroud notes that epistemic partiality—understood as the "unjustified departure from epistemic objectivity"—cuts against purist epistemological standards and norms, such as believing in accordance with the evidence.[19] Friendship,

[12] Goldberg, "Against Epistemic Partiality," 2224-2225.
[13] Stroud, "Epistemic Partiality," 505.
[14] *Ibid.*, 506.
[15] *Ibid.*
[16] *Ibid.*, 507.
[17] *Ibid.*
[18] *Ibid.*, 513.
[19] *Ibid.*, 518.

therefore, opposes mainstream accounts of epistemic rationality.[20] But how should we understand this tension?

Stroud gives three possible explanations for the tension between epistemic partiality in friendship and analytic epistemology. While she does not argue we should prefer any of them, I argue later that we should prefer what I call the radical response. First, friendship may simply require epistemic irrationality (or at least something that epistemic theories of rationality categorize as irrational).[21] According to this response, friendship and epistemic rationality stand (or *can* stand) in irreconcilable conflict with one another, such that we must choose one over the other. If epistemic rationality precludes constitutive elements of the good life, e.g., friendship and, as I'll argue, romantic love, then *so much the worse* for epistemic rationality. We have reasons to prefer elements of the good life over epistemic standards of epistemology. Call this the *radical response*.

Next, Stroud suggests, following Henry Sidgwick, that friendship and rationality might be incommensurate values: "There is what you ought to believe from an epistemic point of view, what you ought to believe as a friend, but no adjudication of those competing claims which gives us what you ought to believe simpliciter."[22] According to this suggestion, when epistemic rationality and friendship demand different beliefs, there is no reason to prefer one to the other. Call this the *incommensurate response.*

Lastly, Stroud suggests that the requirements of the good life are a constraint on epistemology, similar to the constraint of ethical theories. Michael Stocker notes that because "modern ethical theories" exclude relationships that are necessary for the good life we have reason to reject such moral theories.[23] Stroud proposes that a

---

[20] Needless to say, Stroud's view has attracted much dissent. Many reject the proposed conflict between the norms of friendship and epistemic norms. Jason Kawall ("Friendship and Epistemic Norms") and, to some extent, Katherine Hawley ("Partiality and Prejudice") argue that the epistemic demands of friendship fall within the bounds of epistemic propriety because friendship's demands are not as strenuous as Stroud supposes. Lindsay Crawford ("Believing the Best") argues that evidentialist responses to epistemic partiality fail. The attitudes constitutive of friendship, she argues, preclude partiality. Friendship cannot generate reasons in the way Stroud argues. Sanford Goldberg argues that value-reflecting reasons make the demands of friendship and epistemic partiality "epistemically innocuous" ("Against," 2225). Value-reflecting reasons are reasons generated from what we value, which is consistent with our total available evidence. I bracket these objections. My goal here is to further this dissent but in the other direction: I argue there is more conflict between epistemic norms and the norms of the good life than Stroud appreciates.

[21] Stroud, "Epistemic Partiality," 520.

[22] Stroud, "Epistemic Partiality," 519.

[23] Michael Stocker, "The Schizophrenia of Modern Moral Theories," *The Journal of Philosophy* 73, 14 (1973): 453-466.

similar constraint obtains for epistemology: "If standard epistemological theories condemn as irrational something that is indispensable for the good life—so that we have compelling reason *not* to comply with the demands of those theories—then perhaps we should question whether those theories offer an adequate account of epistemic rationality after all."[24] Call this the *constraint response*.

Consider now the limits of Stroud's view: friendship and epistemic partiality does not require we form false beliefs. She writes: "What seems to be characteristic of the good friend is not a stubborn denial of obvious incontrovertible facts about [one's] friend but something more subtle."[25] I argue in the next sections that this is not right; differential epistemic practices may sometimes require forming false beliefs. We might be required to believe falsely in order to love romantic partners for the right reasons or to avoid a doxastic wronging. If this is the case, then romantic relationships sometimes pose epistemic demands that Stroud neglects. I argue, moreover, that if there is a demand for false belief, then we should reject Stroud's constraint response: the epistemic costs—doing away with the truth condition—are too great for any plausible epistemology. Love, in short, requires irrationality: we need, at least, conceptual space for this possibility. I argue that the radical response best affords this space. For now, I turn to the demand to believe falsely within the context of romantic relationships.

4. Love Requires False Beliefs… Sometimes

While romantic relationships may demand epistemic partiality—which does not necessarily entail a demand to believe falsely—I argue that romantic love sometimes poses the further demand to believe falsely. Let's call this the *false belief requirement*. The false belief requirement violates a more fundamental epistemic norm than Stroud's epistemic partiality: the truth condition. We will see how this poses a problem for reconciling norms of the good life and purist epistemology. To substantiate the false belief requirement, I turn to Neil Delaney.[26] I first show that love demands loving for the right reasons; I then argue that this demand sometimes poses the false belief requirement.

Delaney argues that lovers desire to be loved for the right reasons. The right reasons are the properties that the person *takes to be central to* her identity. Delaney writes: "A person A wants a romantic partner B to love him for properties that A takes to be central to his self-conception. Not necessarily all of the properties,

[24] Stroud, "Epistemic Partiality," 522.

[25] *Ibid.*, 506.

[26] Neil Delaney, "Romantic Love and Loving Commitment: Articulating a Modern Ideal," *American Philosophical Quarterly* 33 (1996): 339-356.

perhaps not even just any plurality; but definitely some."[27] If I, for example, deeply identity as a philosopher, my partner must appreciate that part of my identity. Delaney thinks my partner does not have to know about philosophy, but rather she must appreciate that I love philosophy and identify as a philosopher. The idea here is that "people generally want those things about themselves that they take to be at the core of their identity to figure as grounds in the attitude their lover takes toward them."[28] Lovers desire to be seen for what they value about themselves.

Delaney points out that we desire to be loved for the right reasons, even when we have a largely inaccurate conception of ourselves. Consider, for instance, the "goodhearted teenage rebel" with a "decent, gentle soul" who sees himself as a "nihilistic desperado."[29] This boy's dramatic conception of himself is not accurate; he does not want to see himself as he really is, which is kind and gentle. This person, according to Delaney, "wants to be loved for what he *takes* to be central to his self-conception."[30] Lovers desire to be loved for the features they take to be important to their identity, even if that involves a *false* conception of who they are.

I argue that the lover needs to love the features central to her beloved's self-conception, even if those features are false. Delaney gives a recipient-oriented account of loving for the right reasons, as opposed to an agent-oriented account. Delaney's account does not necessarily imply that the lover must love the beloved for the right reasons, merely that the beloved desires that she be loved for these reasons. If the lover is not obligated to love for the right reasons, then the lover can ignore the beloved's desire to be loved for the right reasons. In order for my argument to work, I need the agent-oriented claim that lovers are obliged to love for the right reasons.

In response, it is in the lover's interest to love the beloved for the right reasons. The reason is that the lover also wants to be loved for the right reasons too. There seems to be a reciprocal nature of loving for the right reasons. If the lover knows that he wants to be loved for the right reasons, it seems clear that he ought to also love his beloved for the right reasons.

Loving for the right reasons, moreover, is beneficial to the lover. Love involves benefitting one's lover by letting her be seen as she desires to be seen. Many argue that a desire to benefit one's beloved is partly constitutive of love.[31] Benefit

[27] *Ibid.*, 343.

[28] *Ibid.*

[29] *Ibid.*

[30] *Ibid.*, 344.

[31] See, for instance, Henry Sidgwick, *The Methods of Ethics* (Indianapolis: Hackett, 1874/1981); Gabriele Taylor, "Love," *Proceedings of the Aristotelian Society* 76, 1 (1976): 147-164; Laurence

here involves bringing about the lover's good within reason. If the lover desires to be psychologically seen for certain things, I think love demands we see her for the things she loves about herself. Love, in short, demands loving for the right reasons.

If there is a demand to love for the right reasons and the right reasons are sometimes false, then love sometime poses the false belief requirement. The idea here is that to love in the way that the beloved desires, to give him what he needs, we must adopt some false beliefs about who he is or what he has done. Thus, the demand to love for the right reasons sometimes poses the false belief requirement. Given the false belief requirement, I argue that we should prefer Stroud's radical response.

Now, there may be some limitation to the false belief requirement. Not every false self-conception needs to be loved—"No honey, you're not the greatest criminal mastermind of the twenty-first century"—especially given that Delaney thinks we need not love every identity-constitutive feature of our beloved. But surely there are *some* instances where love makes the false belief requirement. One plausible criterion is that the lover will not obtain psychological fulfillment without being loved for that reason: "[I]f in fact there is a significant disparity, from the point of view of psychological fulfillment they want to think of their lover as valuing them for the same sorts of reasons that they themselves do."[32] Another plausible set of cases where love may require false belief are cases of striving, where we prematurely see ourselves as what we desire to become. Such cases may require taking on dubious epistemic practices. I say more about *striving cases* momentarily, but I first turn to a couple of objections.

Consider another problem. Delaney suspects that discrepancy between the lover's self-conception and who she actually is reveals a defective love. He writes: "I suspect that in cases of radical disparity between who you take yourself to be and who you are fulfilling romantic relationships are effectively precluded."[33] The idea here is that one cannot have fulfilling romantic love and self-discrepancy: it will be in some way deficient. If this kind of love is defective, then the tension between rationality and love is in some sense expected because the love is bad to begin with.

Thomas, "Reasons for Loving," in *The Philosophy of (Erotic) Love*, eds. Robert Solomon and Kathleen Higgins (Lawrence: University of Kansas Press, 1991), 467-476; Harry Frankfurt, "Some Thoughts about Caring," *Ethical Perspectives* 5 (1998): 3-14; Robert Nozick, *The Examined Life* (New York: Simon & Schuster, 1989); and John Rawls, *A Theory of Justice* (Cambridge, Mass: Harvard University Press, 1971).

[32] Delaney, "Loving Commitment," 344.

[33] *Ibid.*, 354, ff 18.

If the love were not defective, then there would be no conflict. Thus, discrepancy between love and epistemic rationality only happens in cases of bad love.[34]

To demonstrate the veridical nature of love, Delaney cites Nathaniel Branden.[35] Branden argues that *psychological visibility* is central to romantic love. Companionship is pleasurable and consequently valuable because it affords the opportunity to feel "seen" by another.[36] We can only, according to Braden, view ourselves conceptually—we know things *about* ourselves—but we need others to view ourselves perceptually, "as concrete objects 'out there.'"[37] Other consciousnesses function like a mirror. Being seen in this way is recognition of personhood. The feeling of being seen is psychological visibility. Romantic love affords a "uniquely powerful" experience of visibility because lovers share a fascination with one another unlike any other relationship.[38]

Psychological visibility requires a veridical self-conception. Branden writes: "If [the other's] view of us is consonant with our deepest vision of who we are (which may be different from whom we profess to be), and if [the other's] view is transmitted by their behavior, we feel perceived, we feel psychologically visible."[39] This conditional seems to assume that we have a largely accurate sense of who we are. To be seen, we must see ourselves as we are.

Branden's account of psychological visibility, as I noted above, spells trouble for my argument. The discrepancy between who we think we are and who we actually are precludes us from becoming psychologically visible.

But I do not think that psychological visibility needs to be totally veridical. Consider cases of striving toward ends. Many of us have experienced the need to change for the better. Whether that be the ambition of accomplishing more than others expect us to or looking at ourselves in the mirror after making a grave mistake, we often desire change. The desire for change sometimes comes with a radical, and often enough premature, change in how we perceive ourselves—I am a good person, or I am a great philosopher. Even if those things are not strictly speaking true at the time, it seems to me that loving for the right reasons means that our lover will see us as what we strive to be.

---

[34] Thanks to Cullin Brown for bringing this objection to my attention.

[35] Nathaniel Branden, "Love and Psychological Visibility," in *Friendship*, ed. Neera Kapur Badhwar (Ithaca: Cornell University Press, 1993), 65-72.

[36] *Ibid*, 67.

[37] *Ibid.*, 69.

[38] *Ibid.*

[39] *Ibid.*, 70.

Sarah Paul and Jennifer Morton argue that whether or not we "believe in others" when they strive toward goals will depend on the relationship we hold with the other person.[40] If this is right, then we ought, it seems, to believe in our lovers when they attempt to accomplish their goals, even if believing that is inconsistent with the evidence. They argue that doubting whether our significant others can achieve their goals can wrong them, in the absence of "significant and specific evidence that this is so."[41] To a certain degree, we must believe in spite of the available evidence to believe our lovers will accomplish their goals. Failing to do so amounts to a wronging. I am arguing here that the goal we strive toward often functions as a reason we desire to be loved for. We want to be seen and valued as a great philosopher, for example, even if that is an end we have yet to achieve.

Part of being psychologically visible, I have argued, sometimes involves being seen as something we are not. That is, psychological visibility does not always strictly speaking require truth. Truth is not required, for example, in striving cases where we desire to premature be seen as what we hope to become. Thus, psychological visibility is not wholly veridical, as Branden and Delaney argue. If Delaney's account of loving for the right reasons is agent-oriented, as opposed to merely recipient-oriented, then the false belief requirement may follow, for example, in striving cases.

Consider yet another problem with my account of Delaney's view: it assumes that we can make ourselves believe a proposition. This view is *doxastic voluntarism*. What I have said supports the following kind of scenario. Albert and Beth are lovers. Albert tells Beth that he is taking up dancing lessons. He *is* a dancer. Beth, who has seen Albert's two left feet in action, knows Albert is a crummy dancer; she also reasonably foresees that Albert will abandon his interest in dancing. She also knows, however, that it is important for Albert to be seen as a dancer, so she makes herself believe that Albert is a dancer, rather than a flippant and temporary dance-enthusiast. That is, she in some way makes herself believe that Albert is a dancer, even though this conclusion is likely false and is formed in spite of the evidence.

Belief formation, according to doxastic voluntarism, is more like doing jumping jacks than a kneejerk reflex. That is, we can come to believe something and stop believing the very same thing at will. This view is scrutinized: many argue that beliefs are at the mercy of evidence. According to these views, we cannot turn beliefs on and off; there is no deliberation about whether to believe whether it is raining or not; we simply come to believe things independently of our control.[42] In response,

---

[40] Sarah Paul and Jennifer Morton, "Believing in Others," *Philosophical Topics* 46 (2018): 75-96

[41] *Ibid.*, 94.

[42] See, for example, Bernard Williams, "Deciding to Believe," in *Language, Belief, and Metaphysics*, eds. Howard Kiefer and Milton Munitz (Albany: SUNY Press, 1970), 95-111; William Alston, "The

there are a growing number of defenders of doxastic voluntarism.[43] I do not have time to contribute to the discussion on doxastic voluntarism, but I rely on the authors cited to demonstrate doxastic voluntarism is plausible. I turn now to a separate but related question.

We might wonder how we can willingly form a false belief. It is a common idea that beliefs track truth, and so we cannot make ourselves believe something false. How is it that I can be required to form a false belief when it is impossible to do so? In response, I return to doxastic voluntarism. As I said, voluntarism says that we have some degree of control over our beliefs. Nishi Shah argues that we have the ability to select different bodies of evidence and slowly influence our beliefs to conform with those various evidence bodies.[44] Matthias Steup argues that our beliefs respond to epistemic reasons, which are to some degree under our control.[45] However it is that we can voluntarily change our beliefs, the false belief requirement will demand it. The false belief requirement might sometimes demand we examine different bodies of evidence or it might demand we change our belief quickly. For now, I return to my discussion of Delaney.

Delaney's claim about loving for the right reasons has an implication for our epistemic norms. I have argued that loving for the right reasons sometimes poses the false belief requirement. That is, we must *believe* certain things about another person that are false. This clearly cuts against epistemic norm of believing truly. Nevertheless, if Delaney is right, then a plausible epistemic demand of romantic relationships is to form beliefs about our beloved's identity, even when those beliefs are not true. Delaney's account can pose the false belief requirement.

Note that this conclusion deviates significantly from Stroud's epistemic partiality. As I've been at pains to show, relationships require more radical deviations

---

Deontological Conception of Epistemic Justification," *Philosophical Perspectives* 2 (1988): 257-299; and Andrei Buckareff, "Acceptance and Deciding to Believe," *Journal of Philosophical Research* 29 (2004): 173-190.

[43] See Carl Ginet, "Deciding to Believe," in *Knowledge, Truth and Duty*, ed. Matthias Steup (Oxford: Oxford University Press, 2001), 63-76; Robert Audi, "Doxastic Voluntarism and the Ethics of Belief," in *Knowledge, Truth and Duty*, ed. Matthias Steup (Oxford: Oxford University Press, 2001), 93-111; Nishi Shah, "Clearing Space for Doxastic Voluntarism," *The Monist* 85 (2002): 436-445; Matthias Steup, "Belief Control and Intentionality," *Synthese* 188 (2012): 145-163; Amy Flowerree, "Agency of Belief and Intention," *Synthese* 194 (2016): 2763-2784; Rima Basu, "Can Beliefs Wrong?," *Philosophical Topics* 46 (2018): 1-17; and Rima Basu and Mark Schroeder, "Doxastic Wronging," in *Pragmatic Encroachment in Epistemology*, eds. Brian Kim and Matthew McGrath (New York: Routledge, 2018), 181-205.

[44] Shah, "Clearing Space," 436-445.

[45] Steup, "Belief Control," 145-163.

from epistemic norms than Stroud's account shows. Recall that her account of epistemic partiality excludes the false belief requirement. Loving for the right reasons points to a much more radical epistemic tension than Stroud's epistemic partiality. None of the mechanisms of epistemic partiality require that we believe falsely about our friend: we may merely require more evidence or interpret character claims charitably. The epistemic requirements of love involve worse epistemic behavior because love can sometimes demand we hold a false belief about our lover. This feature of romantic love, i.e., the more radical epistemic responsibility we hold to our lover, gives us reason to think that romance sometimes *just* requires irrationality, or behaviors that analytic epistemology would categorize as irrational.

To further my argument, I turn now to another facet of loving relationships: the demand not to wrong one another. Specifically, I examine doxastic wronging in the context of loving relations. Since true beliefs can sometimes wrong, lovers may need adopt false beliefs in order to avoid doxastically wronging one another. This facet of love, I conclude, can pose the false belief requirement.

5. Doxastic Wronging and Romantic Love

Doxastic wronging reveals the demand to believe falsely within the context of romantic relationships. This section has three movements. I first argue that romantic partners are especially susceptible to doxastic wronging. Given the intimacy and vulnerability required for romantic relations, lovers stand in a very special position with one another—one in which they are especially prone to doxastic wronging. Second, following Rima Basu,[46] I argue that true beliefs can also wrong. True beliefs that wrong, I argue lastly, reveal a demand for lovers to sometimes believe in spite of the available evidence or even falsely. Lovers hold special responsibilities not to doxastically wrong one another, even with true beliefs. Love's demand not to doxastically wrong one's lover can sometimes pose the false belief requirement.

According to Rima Basu and Mark Schroeder, the content of some beliefs can constitute a moral wronging.[47] It is not the manner in which the belief was formed, e.g., whether the belief was formed on the basis of some morally problematic bias or faulty evidence, nor is it the actions and dispositions that follow from the belief. Doxastic wrongings occur at beliefs themselves, not "upstream" or "downstream" from the belief.

[46] Rima Basu, "What We Epistemically Owe to Each Other," *Philosophical Studies* 176 (2019): 915-931.

[47] Basu and Schroeder, "Doxastic Wronging," Basu, "What We Owe;" Mark Schroeder, "When Beliefs Wrong," *Philosophical Topics* 46 (2018): 115-127.

Basu and Schroeder identify three conditions of a doxastic wronging, which I call the *directed condition*, the *belief itself condition*, and the *content condition*.[48] The directed condition says that a belief wrongs a particular person; it is not just wrong in general. The belief itself condition is that holding the belief—as opposed to how the belief was formed or the actions and dispositions that follow from the belief—is what wrongs. And, lastly, the content condition is that the content of the belief wrongs. Doxastic wrongings obtain in virtue of what is believed.

To illustrate doxastic wronging within the context of love relations, I turn to a touchstone case that supports doxastic wronging:

> *Wounded By Belief.* Suppose that Mark has an alcohol problem and has been sober for eight months. Tonight, there's a departmental colloquium for a visiting speaker, and throughout the reception, he withstands the temptation to have a drink. But, when he gets home his partner, Maria, smells the wine that the speaker spilled on his sleeve, and Mark can tell from the way Maria looks at him that she thinks he's fallen off the wagon. Although the evidence suggests that Mark has fallen off the wagon, would it be unreasonable for Mark to seek an apology for what Maria believes of him?[49]

Something is the matter with Maria's belief. It fails to appreciate Mark's determination to resist drinking; what is an accomplishment to Mark is a sign of defeat for Maria. Basu and Schroeder conclude that Maria has wronged Mark with the content of her belief.[50]

Does it matter that Maria is Mark's partner? I believe so. If Maria were a stranger on a train, for instance, it doesn't seem to matter quite as much that she believes that Mark drank.[51] Mark and Maria, we might assume, have been working

---

[48] Basu and Schroeder, "Doxastic Wronging," 181-205.

[49] Basu, "What We Owe," 917.

[50] Here one might wonder whether if it is the belief itself that wrongs or Maria's expression that follows from the belief. Consider another case from Basu ("What We Owe") to invoke the belief contents can wrong: *Racist Hermit*. "One day a racist hermit finds a newspaper highlighting the academic success of Sanjeev. The hermit concludes that Sanjeev smells of curry. Suppose that Sanjeev has recently made curry—rendering the hermit's belief true." There's no risk of the hermit meeting Sanjeev, nor contributing to any racist structures, yet his belief still seems wrong. One explanation is that the hermit wrongs Sanjeev with the content of his belief.

[51] What's the difference between the racist hermit's belief and the stranger's belief? One is a doxastic wronging, while the other is not. The racist hermit's belief is a doxastic wronging because it is based on a generalization. The stranger's belief that Mark has had a drink does not appeal to a generalization. Rather, her belief is specific to Mark, yet the stranger does not know that Mark has struggled for sobriety. In order to doxastically wrong Mark, I think one must know about his alcoholism. She does not understand the significance of her belief.

on Mark's drinking for a long time. Maria understands how hard Mark has worked and, therefore, should view the night as a success. If Maria is a stranger, she will not know to view the dinner party as a success for Mark's drinking. If Maria infringes on an epistemic duty to Mark in virtue of being Mark's lover, then it seems that lovers must believe rightly about one another even when the evidence suggests otherwise—lovers, that is, possess epistemic duties *because* they are lovers.

Lovers stand in a privileged position with one another—they know things not usually shared with other people. This special position makes us especially vulnerable to doxastic wronging. Perhaps Mark's colleagues do not know he is a recovering alcoholic. Maria, perhaps, is the only one who can wrong him in this way. She is, perhaps, the only one with the responsibility to trust him about his recovery.

We might imagine that Mark trusts Maria more than other people. Given the proximity of their relation, it matters to Mark what Maria thinks about him, and vice versa. I think that Mark's potential expectation for Maria to believe well of him makes him especially vulnerable to doxastic wronging. In many cases, we would not care what a stranger believes of us, where we would care what our lover believes of us.

As lovers, we stand in a special position to doxastically wrong. We know more about our lovers than anyone else, so we can more easily wrong them with our beliefs. Though strangers can doxastically wrong one another—like racist beliefs—lovers have made themselves vulnerable to one another; lovers know one another's weaknesses, just like Maria knows Mark's secret alcoholism. It follows that her believing Mark drank is a doxastic wronging, whereas if a stranger formed such a belief, no apology would be necessary. This is because the stranger does not have the larger context of Mark's alcohol recovery.

So far, I have argued that lovers stand in an especially vulnerable position with one another with regard to doxastic wronging. I now turn to the question of whether beliefs must be false to wrong. There is nothing about a belief's being true that precludes it from wronging—even true beliefs, therefore, can doxastically wrong others.[52] Given that true beliefs can wrong, it follows that sometimes we should form

[52] I follow Basu ("What We Owe") in making this point. It is, however, controversial. Mark Schroeder ("When Beliefs Wrong") argues that doxastic wrongings only occur when the belief in question is false. This poses some issue with my argument that true beliefs can doxastically wrong. He distinguishes between subjective and objective wrongings. For example, suppose Richard is behind one of three closet doors and Gretchen fires a gun at one of these doors, missing Richard. We would be inclined to say that Gretchen's action is both objectively and subjectively wrong: objectively wrong because she poses a needless risk on Richard and subjectively wrong because she did not have sufficient evidence to believe that Richard was not behind the door she shot.

false beliefs about other people. Given that doxastic wronging does not occur in virtue of problematic belief formation processes, beliefs can wrong even if they are formed consistently with the available evidence. Thus, beliefs can wrong even when they are the product of reasonable epistemic processes and are true: none of the essential conditions of doxastic wronging precludes this possibility.

Basu argues that even true beliefs can wrong.[53] Included in this responsibility, therefore, is the demand to refrain from holding some true beliefs. I argue, momentarily, that lovers may sometimes be required to refrain from forming true belief and may be required to adopt false beliefs. But first, consider a case meant to show true beliefs can wrong:

> *Racist Hermit*. Suppose a racist hermit in the woods discovers trash containing an alumni newsletter from Sanjeev's university, which includes Sanjeev's photo. The hermit immediately concludes that the pictured person—Sanjeev—smells of curry. Suppose also that Sanjeev happens to have recently made curry, so in this instance the hermit's belief is true—Sanjeev *does* smell of curry. Has the hermit wronged Sanjeev?[54]

Here is a proposed example of a belief that is true but nevertheless wrongs another person. The hermit, in order to avoid the wronging, must either withhold belief or believe falsely. If asked, the hermit ought to respond that Sanjeev does not smell of curry or that he has withheld belief. But the way we form beliefs is often automatic; beliefs simply appear. If the hermit can't help forming a belief in this case, then it seems that he must form the false belief that Sanjeev does not smell of curry. That is, the hermit is obligated, in order to avoid the doxastic wronging, to form a belief that is false, namely that Sanjeev does not smell of curry.[55]

Objective and subjective wrongings can however come apart. Beliefs too can objectively and subjectively wrong, and they can come apart for beliefs too. For example, according to Schroeder, if a belief would wrong another person, that counts as a reason not to adopt that belief. Suppose that a belief is true, and someone adopts it without having further evidence. That person would have subjectively wronged whoever he believes about, but he would not have objectively wronged him since the belief is true. He subjectively wrongs him because of lack of evidence. If Schroeder is right, then my claim that true beliefs can wrong is not right; however, true beliefs, even though they cannot objectively wrong, can subjectively wrong on this account. Thus, it might be that lovers must refrain from believing when their beliefs would subjectively wrong—and, it is possible, the false belief requirement may still stand. For now, I continue operating under the assumption that true beliefs can wrong.

[53] Basu, "What We Owe," 915-931.

[54] *Ibid.*, 919.

[55] There is again the question of whether doxastic voluntarism is true. I point to the authors cited above for a defense of voluntarism.

But we might wonder what the wronging is here. Perhaps the wronging here is not the belief that Sanjeev smells of curry, rather the wronging is believing that Sanjeev and others of his ethnic group are inferior *because* they smell of curry. The hermit can believe that Sanjeev smells of curry without also believing that Sanjeev and members of his ethnic group are inferior and vice versa. If this is right, then this is not an example of a true belief that wrongs, since the true belief is does not by itself wrong. The inference, rather, to the false belief that Sanjeev and members of his ethnic group are inferior is the wronging.

In response, Basu finds it intuitive that the belief that Sanjeev smells of curry is a wronging. She thinks that there is some harm involves in the hermit's belief: "the harm is a *relational* harm: the hermit fails to relate as he ought."[56] This seems right, but there is more to be said here. Believing on the basis of stereotypes, I think, is the relational failure here. The reason is that stereotypes fail to appreciate an individual as distinct from her reference class.[57] The hermit does not consider the possibility that Sanjeev has distinct features from his reference class: perhaps he does not smell of curry because he does not like curry. To relate to Sanjeev in the right way, the hermit must consider him as an individual. Failing to see Sanjeev as an individual is what makes the hermit's belief wrong. Thus, the hermit's true belief still seems to wrong Sanjeev.

That true beliefs can wrong is an important claim for my argument. If true belief can wrong, then we must either withhold belief or believe falsely, else we commit a wrongdoing. Lovers, who as I have argued have special obligation not to doxastically wrong, will sometimes have to either refrain from believing or believe falsely.

Suppose for example that Maria's belief that Mark has had alcohol is true. Even if that belief is true, Maria's belief may still wrong Mark. Maria fails to take seriously the possibility that the wine was there by accident, even if Mark did drink. I suggest that the demand not to doxastically wrong sometimes requires believing falsely. Perhaps Maria, for example, should still believe that Mark had not had anything to drink, even when the belief is true. If Maria has to form a belief, then it seems like, given her proximity to Mark, she ought to default to the belief that he has resisted alcohol. Perhaps Maria's belief constitutes a wronging until Mark comes clean about his drinking.

---

[56] Basu, "What We Owe," 919. Basu's emphasis.

[57] For more about the moral wrongings that stem from failing to distinguish individuals from their reference class, see J. Spencer Atkins, "Moral Encroachment, Wokeness, and the Epistemology of Holding," forthcoming in *Episteme*.

Now I argue that lovers possess a special obligation not to wrong one another. I think that the demand not to doxastically wrong our lover comes from a common feature of love: the desire to benefit or bring about the good of the beloved. Many authors identify love as a two-pronged desire: the desire for union with the beloved and the desire for the lover's good.[58] To wrong another with one's beliefs is largely inconsistent with the desire to bring about the good of the beloved. We might also think, moreover, that doxastic wronging inhibits union with the beloved, as wronging another person often drives her away. The desires of love—to benefit and share union with—seem to point to the demand, even the self-imposed demand, to refrain from doxastic wronging. It, at the very least, points to the conclusion that to act consistently with love, one would refrain from doxastic wrongings.

Love, given the intimacy and vulnerability of the romantic relation, must demand lovers not wrong one another, including with belief. The plausibility of doxastic wronging, therefore, points to the conclusion that love can sometimes pose the false belief requirement, at least until there is more available evidence. I now draw out an implication of this conclusion: that the best explanation for the false belief requirement is the radical response—that love and epistemic rationality, specifically believing truly, are sometimes mutually exclusive.

6. In Favor of the Radical Response

Recall that Stroud gives three possible interpretations of the relationship between for partiality and epistemic norms: *radical response*, *incommensurate response,* and *constraint response*. She thinks that each of these responses can account for the relation between partiality and epistemic norms. I argue now that the *radical response* is the best explanation of epistemic norms and love's demand to believe falsely sometimes. The false belief requirement poses too strong of a constraint on purist epistemology. That is, adjusting purist epistemology to meet love's demand to sometimes believe falsely jeopardizes a foundational assumption of epistemology, to believe truly. Love and epistemic rationality stand opposed to one another.

The *constraint response* can reasonably explain Stroud's epistemic partiality in friendship but not the false belief requirement. According to this response, the demand for epistemic partiality in friendship requires that we take epistemic

[58] J. David Velleman ("Love as a Moral Emotion," *Ethics* 109 (1999): 338-374) identifies this view of love as the predominant view among analytic philosophers. Given the wide acceptance of this view of love, I limit my discussion to this view. Authors that hold some version of this view are Henry Sidgwick (*Methods of Ethics*), Gabriele Taylor ("Love"), Laurence Thomas ("Reasons for Loving"), Harry Frankfurt ("Some Thoughts"), Robert Nozick (*The Examined Life*), and John Rawls (*A Theory of Justice*).

rationality back to the drawing board. We need, more specifically, to come up with an account of epistemic rationality that leaves room for partiality toward friends. Such an account of rationality could be plausibly consistent with many other epistemic norms.[59] *Constraint response*, however, fails to account for love's demand to adopt false beliefs in some circumstances. Recall that truth is an indispensable concept for analytic epistemology, as I tried to show in the first section of the paper. I have argued, however, that love sometimes requires that we hold false beliefs. Coming up with an account of rationality that is compatible with holding false beliefs strikes me as a losing battle: an account of rationality that does away with truth is too strong of a constraint. There are some central epistemic norms that any plausible account of rationality must respect—one of which is true belief.

Moreover, I do not think that the *incommensurate response* is helpful either. Recall that this response says that the value of rationality and the values of friendship and love are neither lesser than, greater than, nor equal to one another; they are just different. We, consequently, have no reason to prefer one to the other, given that they do not stand in any hierarchic relation to one another. I do not think these values are incommensurate, but even if they are incommensurate, we often treat them as though they are not mutually exclusive. It's intuitive to think that we can have both for the most part. If these values were truly incommensurate, then we would, I believe, often expect love and rationality to come apart, and not care if they did. But we do care and we do not often expect them to come apart. Perhaps this offers some reason to think love and rationality are commensurate values.

The *radical response* is the best explanation of the false belief requirement—there sometimes *just is* an unavoidable conflict between two competing values. To illuminate this conclusion, consider Tamar Gendler's observation about rationality and racism: "Living in a society structured by race appears to make it impossible to be both rational and equitable."[60] Gendler argues that rationality will often require us to form belief that are consistent with various base rates, where base rates are understood as statistical trends within a given context. Such trends, however, can be racist. For instance, suppose that the swanky DC night club, the Cosmos Club, has nearly all black employees and nearly all white club members. A person looking for an employee, where employees and club members both wear tuxedoes, would be epistemically rational to believe that some particular black person is an employee, given the base rate at the Cosmos Club. In a similar situation, a club member mistook

[59] Kawall ("Friendship and Epistemic Norms") argues for a similar point. That is, he argues that epistemic partiality is not precluded for purist epistemic standards. The idea here is that purist epistemic norms can be adjusted such as to include epistemic partiality.

[60] Tamar Gendler, "On the Epistemic Cost of Implicit Bias," *Philosophical Studies* 156 (2011): 57.

black historian John Hope Franklin for an employee. Franklin writes that he felt wronged by the club member. The club member's belief, however, is formed in accordance with the evidence and, by many epistemic standards, quite rational. It is also racist. Gendler concludes that there is irreconcilable conflict between epistemic rationality and moral norms, specifically not holding racist beliefs. In a similar vein as Gendler, I have said that there is a conflict between epistemic rationality and the norms of love. Epistemic rationality would require us to be bad lovers in some circumstances. Believing truly makes for bad love. The sad conclusion is that sometimes we are forced to choose between being rational or being good lovers; we cannot always have both.

Consider a belief objection. Perhaps the beliefs we form in loving relationships involve non-propositional knowledge, knowledge by acquaintance. It might be that loving for the right reasons is, for instance, a kind of non-propositional activity. The reasons I love my partner are not true and false propositions, rather they are experiences and sensations, e.g., the experience of my lover's perfume or the way she looks when she gets mad. Bertrand Russell[61] argues that knowledge by acquaintance does not involve inference, judgement, or thought. This kind of knowledge is not subject to bivalence, as propositional knowledge is. Knowledge—as defined by many analyses of knowledge—however deals with propositional knowledge; the domain of modern analytic epistemology overwhelmingly deals propositional knowledge—knowledge *that*—not knowledge by acquaintance.[62] If this is the case, then it seems that modern analyses of knowledge will sidestep the problems I have laid out here because knowledge of one's lover is non-propositional and, consequently, outside of the scope of the epistemology I have critiqued.

In response, loving relationships likely have some elements of knowledge by acquaintance. I may enjoy—beyond the mere fact *that* my lover wears perfume—the experience of the smell of my lover's perfume. But this does not mean that every reason that we love our lover is non-propositional. Loving for the right reasons *is* often propositional because we can, as I argued, believe various false propositions about the lover, e.g., that he is a "nihilistic desperado." Thus, this objection does not capture all of the logical space. There are some instances where we must assume false propositions (that go beyond mere knowledge by acquaintance) to love our lover for the right reasons.

[61] Bertrand Russell, *The Problems of Philosophy* (London: Williams and Norgate, 1912).
[62] Ichikawa and Steup, "The Analysis of Knowledge."

7. Conclusion

Love sometimes poses the false belief requirement. Love sometimes requires *bad epistemic agency*. Specifically, I argued that the epistemic norm of believing truly sometimes stands opposed to both the demand to love for the right reasons and the demand to avoid doxastic wronging. Stroud's constraint response—that the epistemic norms must subordinate themselves to the norms of the good life—goes too far. Epistemology cannot give up the truth condition. I proposed that this conflict is best explained by the radical response: that we must sometimes be irrational if we are to love another person. Prioritizing epistemic rationality over the demands of our lover means forsaking the demands of love. Forsaking such demands, I think, inhibits the good life. If epistemic demands inhibit the demands of love—and consequently the good life—then so much the worse for epistemic demands.

METHODOLOGICAL NATURALISM AND REFLEXIVITY REQUIREMENT

Hamed BIKARAAN-BEHESHT

ABSTRACT: Methodological naturalists regard scientific method as the only effective way of acquiring knowledge. Quite the contrary, traditional analytic philosophers reject employing scientific method in philosophy as illegitimate unless it is justified by the traditional methods. One of their attacks on methodological naturalism is the objection that it is either incoherent or viciously circular: any argument that may be offered for methodological naturalism either employs a priori methods or involves a vicious circle that ensues from employing the very method that the argument is aimed to show its credentials. The charge of circularity has also been brought against the naturalistic arguments for specific scientific methods; like the inductive argument for induction and the abductive argument for the inference to the best explanation. In this paper, I respond to the charge of circularity using a meta-methodological rule that I call 'reflexivity requirement.' Giving two examples of philosophical works, I illustrate how the requirement has already been considered to be necessary for self-referential theories. At the end, I put forward a meta-philosophical explanation of the naturalism-traditionalism debate over the legitimate method of philosophy.

KEYWORDS: methodological naturalism, reflexivity, scientific method, armchair philosophy, a posteriori argument

1. Introduction: Methodological Naturalism

Methodological naturalism is frequently defined as a commitment to employ scientific method – or more generally, empirical methods – in philosophy as well as science. It means that philosophers, like scientists, should exclusively make use of scientific method to address philosophical problems, since from the naturalists' point of view, it is the only effective way of acquiring knowledge of the world. Such a reading of methodological naturalism has already appeared in many works concerning naturalism. For example, in *Encyclopedia of Philosophy*, Keith Campbell views the methodological aspect of naturalism in this way:

> […] naturalism is sometimes regarded as a rule of method rather than a metaphysical doctrine. There is a natural method of inquiry, which consists in setting out to explain and understand the world by finding the natural causal processes by which natural objects come into being, produce their effects, and pass

> away. *All genuine knowledge* is of this natural, experimental kind; […].[1]

Jaegwon Kim also defines 'the epistemological thesis' of naturalism as the idea that "*all knowledge that we can acquire* is acquirable *only* through the application of scientific method."[2] As another instance, Mario De Caro and David Macarthur characterize a "methodological (or epistemological) scientific naturalist" as someone who believes "it is *only* by following the methods of the natural sciences – or, at a minimum, the empirical methods of a posteriori inquiry – that one arrives at genuine knowledge."[3] Likewise, Michael Devitt defines naturalism as the thesis that "there is *only one way* of knowing, the empirical way that is the basis of science (whatever that way may be)."[4] And many anti-naturalists understand methodological naturalism in the same way.[5] So, I presume the following thesis to be the main claim of methodological naturalists:

> (**MN**): Scientific method is the only route to knowledge (in all areas including philosophy).

This line of thought constitutes an integral part of *philosophical* naturalism in general.[6] But, what is the reason behind the naturalists' strong belief in the view so

---

[1] Keith Campbell, "Naturalism," *Encyclopedia Of Philosophy*, Vol. 6, ed. Donald M. Borchert (2005), 492 (my italics).

[2] Jaegwon Kim, "The American Origins of Philosophical Naturalism," *Journal of Philosophical Research* 28, Supplement (2003): 83-98, 88 (my italics).

[3] Mario De Caro and David Macarthur, "Introduction: The Nature of Naturalism," in *Naturalism in Question*, ed. Mario De Caro and David Macarthur (Cambridge: Harvard University Press, 2004), 1-17, 7 (original italics).

[4] Michael Devitt, *Coming to Our Senses: A Naturalistic Program for Semantic Localism* (Cambridge: Cambridge University Press, 1996), 2; Michael Devitt, "Naturalism and the a Priori," *Philosophical Studies* 92 (1998): 45-65, 45 (my italics). See also Devitt, "Naturalism and the a Priori," 46-47.

[5] See, e.g., Robert Audi, "Philosophical Naturalism at the Turn of the Century," *Journal of Philosophical Research* 25 (2000): 27-45, 41; Laurence BonJour, "Against Materialism," in *The Waning of Materialism*, ed. Robert C. Koons and George Bealer (Oxford: Oxford University Press, 2010), 3-23, 7. For some slightly different (less extreme) definitions of methodological naturalism, see Kim Sterelny, *The Representational Theory of Mind: An Introduction* (Oxford: Basil Blackwell, 1990), xi; Tim Lewens, "A Surfeit of Naturalism," *Metaphilosophy* 43, 1-2 (2012): 46-57; De Caro and Macarthur, "Introduction;" Philip Kitcher, "Giving Darwin His Due," in *The Cambridge Companion to Darwin*, ed. Jonathan Hodge and Gregory Radick (Cambridge: Cambridge University Press, 2009), 455-476.

[6] Wilmon Henry Sheldon, "Critique of Naturalism," *The Journal of Philosophy* 42, 10 (1945): 253-270; Alexander Rosenberg, "A Field Guide to Recent Species of Naturalism," *The British Journal for the Philosophy of Science* 47, 1 (1996): 1-29; Alexander Rosenberg, "Disenchanted Naturalism," in *Contemporary Philosophical Naturalism and Its Implications*, ed. Bana Bashour and Hans D.

central to their position? That is, what is their justification for (MN)? Finding the answer to this question will become more important when it is realized that many works in philosophy draw heavily on the thesis, and there is a great controversy over it.

I start finding an answer to the question with scrutinizing (MN). (MN) has an implicit negative component: it implies that traditional methods of *armchair philosophy* or *first philosophy* fall short of directing us toward knowledge, and, as a result, they should be abandoned.[7] The main reason naturalists have for this claim is not that traditional methods are illegitimate – that is, resulting in false beliefs[8] – but it is the supposed unfruitfulness and unproductiveness of those methods.[9] Philosophy, they claim, has been extremely unsuccessful and it has been due to the infertile methods of first philosophy. On the other hand, for them, most (or even all) the knowledge we have about the world has been acquired using scientific method, that is, science achieved its success via its method.[10] Then, scientific method can be vindicated according to its past successes. If one regards such a reason as an argument for methodological naturalism – as many naturalists claim[11] – it will be an *a posteriori argument*. It is because in that argument, some evidence from history of science and history of philosophy is provided for defending (MN).

(MN) has come under two criticisms. Firstly, it is not clear what is meant by 'scientific method.' Actually, it is hard to determine a necessary and sufficient condition a kind of method should meet for being scientific. I am not going to tackle with this problem, and it does not affect what I am arguing for here. The only thing I need here is that the a posteriori way of justifying (MN) can be considered to be scientific in a sense. Secondly, contrary to what is claimed by naturalists, many anti-

---

Muller (New York: Routledge, 2014), 17-36; Kim, "American Origins;" Audi, "Philosophical Naturalism at the Turn of the Century"; Timothy Williamson, "What Is Naturalism?," in *Philosophical Methodology: The Armchair or the Laboratory?*, ed. Matthew C. Haug (New York: Routledge, 2014), 29-31; Daniel Stoljar, *Physicalism* (London: Routledge, 2010), 11.

[7] Willard Van Orman Quine, *Theories and Things* (Cambridge, MA: Harvard University Press, 1981), 72.

[8] It seems that some naturalists even tend to make that claim. See, e.g., Daniel C. Dennett, "Current Issues in the Philosophy of Mind," *American Philosophical Quarterly* 15, 4 (1978): 249-261, 250.

[9] See John Dewey, *The Influence of Darwin on Philosophy and Other Essays in Contemporary Thought* (New York: Henry Holt and Company, 1910), 17-18.

[10] See, e.g., Alexander Rosenberg, "Can Naturalism Save the Humanities?," in *Philosophical Methodology: The Armchair or the Laboratory?*, ed. Matthew C Haug (New York: Routledge, 2014), 39-42.

[11] See Ronald N. Giere, "Modest Evolutionary Naturalism," *Biological Theory* 1, 1 (2006): 52-60, 53; Hilary Kornblith, "Naturalism: Both Metaphysical and Epistemological," *Midwest Studies in Philosophy* 19, 1 (1994): 39-52, 49.

naturalists believe that the method of philosophy must be certain and scientific method whose infallibility is not beyond doubt is not legitimate in philosophy. (I will call these anti-naturalists *methodological traditionalists* – traditionalists for short.[12]) Hence, not only do traditionalists want to undermine (MN), but also they insist on a more extreme view about the illegitimacy of using scientific method in philosophy such that they even reject the more moderate versions of naturalism. However, one of the most important counterarguments of traditionalists against methodological naturalism has been posed against (MN). As explained in the next section, via that argument, traditionalists aim to block *any* kind of argument naturalists *may* adduce for methodological naturalism. If successful, such an argument deprives naturalists of having any argument for their position.

2. *A Priori Argument* against Methodological Naturalism: The Charge of Circularity

The simple form of the counterargument against methodological naturalism proceeds with a dilemma: naturalists either justify (MN) *a priori* or justify it *a posteriori*. According to (MN) itself, a priori justification is not allowed. Then, the first horn of the dilemma leads naturalists to some kind of incoherence.[13] Therefore, naturalists have no choice rather than the second horn, i.e. to go on an a posteriori argument. As described in the previous section, this is their actual choice. But, is a posteriori justification legitimate in philosophy? As said before, according to *standard* armchair philosophy, the answer is negative unless there is some a priori argument for legitimacy of a posteriori justification. Now, if naturalists say that they consider a posteriori justifications to be as legitimate as a priori ones, then,

[12] Siegel ("Naturalized Epistemology and 'First Philosophy'," *Metaphilosophy* 26, 1-2 (1995): 46-62, esp. 48-49; "Empirical Psychology, Naturalized Epistemology, and First Philosophy," *Philosophy of Science* 51, 4 (1984): 667-676, esp. 671-673) rejects such a characterization of traditionalism. However, even if he is right, his idea of 'moderate first philosophy' – including such ideas as 'justification from without' and 'nonempirical theory of justification' – may be used here to characterize traditionalism without having a significant effect on my discussion. It is because any idea that seeks an 'extrascientific basis' for science will not be acceptable from a naturalistic point of view (see Giere, "Modest Evolutionary," 53), since naturalists are in conflict with traditionalists over "the (il)legitimacy of appealing to the results of scientific inquiry in order to establish the epistemic credentials of (the results of) that selfsame inquiry" (Siegel, "Naturalized Epistemology," 49 (fn. 45)). Then, whichever way we understand traditionalism (or first philosophy), it is no less than a struggle to analyze science 'from without' to avoid "a seemingly vicious circularity problem" which naturalists deny (see Siegel, "Naturalized Epistemology," 52. See also Paul A. Roth, "Siegel on Naturalized Epistemology and Natural Science," *Philosophy of Science* 50, 3 (1983): 482-493).

[13] See Giere, "Modest Evolutionary," 53.

traditionalists claim, they indulge themselves in a vicious circle, because they suppose the claim they are supposed to prove. So, in either horn, naturalists are unable to justify their position, and then, they are unjustified. The situation can be called a 'naturalist's dilemma.'[14]

The dilemma has been touched on by some anti-naturalists; among them are Robert Audi and Harvey Siegel. Audi says that if we characterize naturalism as prescribing 'certain basic methods of inquiry,' then, compared with other methods, it considers those methods privileged and as a result, naturalism is subject to 'actual-method chauvinism' – the claim that the only legitimate methods are some specific actual methods of science. Then, he continues:

> I find it hard to see how to justify such privilege except possibly on a priori grounds. […] unless we are to acquiesce in skepticism, we must apparently choose between *a priori justification* on the one hand and, on the other, either forgoing justification or countenancing *circular justification*.[15]

The dilemma Audi describes is exactly the naturalist's dilemma.

Siegel, when criticizing Gibson's defense of Quine's naturalism, makes use of the same tactic to disarm naturalists: "if the premises of any such [naturalistic] argument [for naturalism] are justified non-naturalistically, then self-referential incoherence threatens; if they are justified naturalistically, they will be unable to play any non-circular (or non-question-begging) role in the justification of naturalism."[16] He makes the same point against Roth's defense of Quine's naturalism, saying that: "This seems to me a very deep problem for the naturalized epistemologist: it must assume the legitimacy of, and strive to achieve, the very sort of justification it seeks to show cannot be had."[17] Naturalists themselves have considered the criticism seriously.[18]

I call the anti-naturalist counterargument 'a priori argument against methodological naturalism' – *a priori argument* for short. For assessing the argument, it is better to formulate it first. If we take (MN) for granted, then we can formulate the argument as follows:

> **Argument (AA)**: a priori argument against methodological naturalism
>
> 1. If (MN) is justified, it is justified either a priori or a posteriori.

[14] This expression has occasionally been used in slightly different meanings. See Audi, "Philosophical Naturalism," 42; Williamson, "What Is Naturalism?," 30.

[15] Audi, "Philosophical Naturalism," 42 (my italics).

[16] Siegel, "Naturalized Epistemology," 59.

[17] See Siegel, "Empirical Psychology," 675-676. See also Paul A. Roth, "Theories of Nature and the Nature of Theories," *Mind* 89, 355 (1980): 431-438; Roth, "Siegel on Naturalized," 485.

[18] See, e.g., David Papineau, *Philosophical Naturalism* (Oxford: Blackwell, 1993), 3-4.

2. A priori justification is not a proper justification.
3. If (MN) is (only) justified a posteriori, then its justification is viciously circular. (**VC**)
4. A viciously circular justification is not a proper justification.

∴ (MN) does not have a proper justification.

(AA) is a dialectical argument, since through it, it is shown that with espousing (MN), naturalists will be deprived of having a justification for their position and so, their position are totally unjustified.

All the premises of (AA) are supposed to be a priori. The premise (1) is based on the fact that every justification is exclusively either a priori or a posteriori. The premise (2) is a logical consequence of (MN) together with the assumption that a justification is proper only if it makes use of a method which can lead us to knowledge. The premise (4) seems to be endorsed by most epistemologists. These premises do not seem challenging. The main challenge of the argument is about the premise (3) – which I named (VC) – and the responses to the argument have mostly been accompanied with rejecting this premise. In the next section, I examine two responses to this argument and to the similar counterargument against the inductive argument for induction.

3. Two Responses to the Charge of Circularity

As I said before, some of the naturalists have considered the a priori argument seriously and responded to it. In this section, I examine two responses to the charge of circularity, one to the a priori argument (AA), and another to the similar objection to the inductive argument for induction.

3.1 Papineau's Response

David Papineau responds to the a priori argument in this way: traditionalists in that argument do not simply rest on a concept of philosophical activity; in fact, they have some serious presuppositions, the most important one is the assumption that knowledge should be certain, that is, it should be acquired via methods which *necessarily* deliver truth.[19] Applying the response to (AA), we can see that Papineau's main target is the premise (VC). According to (VC), if we employ an a posteriori method to justify (MN), we will be indulged in a vicious circle, since a posteriori methods are not legitimate in themselves, and they can only be employed with taking an assumption like (MN) for granted. So, traditionalists have the supposition that a posteriori methods are illegitimate until they are justified via a

[19] Papineau, *Philosophical Naturalism*, 4.

priori methods. But, Papineau rejects this supposition as unfounded. According to him, "[k]nowledge […] is the state that we need to get into if we are to succeed in avoiding error."[20] So, a belief-forming process (here, a method) to be "an effective means of avoiding error," should only be reliable in actual world, and not necessarily certain – that is, it needn't be "*impossible* for a given belief-forming process to produce a false belief."[21] Then, under this definition of knowledge, scientific methods should be regarded as legitimate in so far as they are reliable.[22]

Put it differently, there is a 'dialectical situation' here: each side of the dispute (naturalism vs. traditionalism) has its own methodology and the proponents of each side think that their own position should be justified via legitimate-according-to-them methods. Now, the important question is that "[…] which philosophical methodology should be used to address *this* issue [itself]?"[23] For Papineau, the response is to follow empirical methodology and the main reason behind his response is that "the onus surely lies with those who want to exclude relevant and well-confirmed empirical claims from philosophical debate to provide some prior rationale for doing so."[24] In fact, we can look at Papineau's opinion as though if two sides of a dispute have different viewpoints about the legitimate methods of inquiry, then none of them is allowed to criticize the other's position using her own method. Rather, if one's method *seems initially good enough*, the burden of proof is on the other side to show that her opponent's method is not legitimate (probably with the aid of a more reliable method).

I think that Papineau clarified the issue in the best way and I am in sympathy with his response to the a priori argument. But, I think, his response should be accompanied by a *meta-philosophical* explanation of the 'dialectical situation' and a suggestion to settle the dispute or, at least, a suggestion for how it could be judged. Therefore, although Papineau's response has many things in common with the proposal I am going to put forward in this paper, my proposal, I think, will be more acceptable since it also contains a meta-philosophical explanation of the situation and a suggestion for settling the dispute.

---

[20] Papineau, *Philosophical Naturalism*, 143.

[21] Papineau, *Philosophical Naturalism,* 144 (original italics).

[22] For more details, see Papineau, *Philosophical Naturalism*, 143-152.

[23] Papineau, *Philosophical Naturalism*, 4 (original italics).

[24] Papineau, *Philosophical Naturalism*, 4. See also *Philosophical Naturalism*, 156-157; Rosenberg, "Can Naturalism Save," 39.

3.2 Baithwaite-van Cleve Response

In addition to the a posteriori argument for methodological naturalism, some of the other *naturalistic arguments* which have been put forward to justify the reliability of specific naturalistic methods have been accused to the charge of circularity as well. Richard Fumerton, for instance, accuses an inductive argument for justifying induction (and any other argument of this sort) of circularity. Based on his analysis, he rejects epistemological externalism.[25] Another example is Richard Boyd's naturalistic argument in defense of scientific realism which is known as *explanationist* (or abductive) *defense of realism* (EDR).[26] The argument makes use of inference to the best explanation (IBE) to show that IBE is a reliable method: the best explanation of the success of IBE in science is that it is reliable. Then, using IBE itself, IBE is reliable, and therefore, what our best scientific theories (which are justified using IBE) tell us is nearby true. This argument has been accused of circularity by Arthur Fine on the ground that it makes use of the very method which the argument is supposed to justify.[27]

Richard Braithwaite and James van Cleve separately tried to respond to the potential-in-their-times (later, an actual) accusation of circularity against a naturalistic inductive argument for the reliability of induction.[28] Their response is simply as follows: an argument can be circular in two different ways: 'premise-circular' and 'rule-circular.' An argument is premise-circular if its result is contained in (at least) one of its premises. This kind of circularity is apparently vicious. But, if

---

[25] Richard A. Fumerton, "Skepticism and Naturalistic Epistemology," *Midwest Studies in Philosophy* 19, 1 (1994): 321-340, 337-338; Richard A. Fumerton, *Metaepistemology and Skepticism* (Maryland: Rowman & Littlefield Publishers, 1995), 177,180. See also Laurence BonJour, *Epistemology: Classic Problems and Contemporary Responses*, 2nd ed. (Maryland: Rowman & Littlefield Publishers, 2010), 64.

[26] See Richard N. Boyd, "On the Current Status of the Issue of Scientific Realism," *Erkenntnis* 19 (1983): 45-90; Arthur Fine, "Unnatural Attitudes: Realist and Instrumentalist Attachments to Science," *Mind* 95, 378 (1986): 149-179; Stathis Psillos, *Scientific Realism: How Science Tracks Truth* (London: Routledge, 1999), chap. 4.

[27] See Arthur Fine, "The Natural Ontological Attitude," in *Scientific Realism*, ed. Jarrett Leplin (Berkeley, CA: University of California Press, 1984), 83-107, esp. 85-86; Fine, "Unnatural Attitudes."

[28] See Richard Bevan Braithwaite, *Scientific Explanation: A Study of the Function of Theory, Probability and Law in Science* (Cambridge: Cambridge University Press, 1953); James van Cleve, "Reliability, Justification, and the Problem of Induction," *Midwest Studies in Philosophy* 9, 1 (1984): 555-567. Psillos (*Scientific Realism*) follows their approach to respond to the accusation of circularity against EDR.

an argument proceeds with the same method which the argument is aimed to justify, it is rule-circular and such circularity is not vicious at all.[29]

Now, if we are to use this strategy to respond to the charge of circularity against the a posteriori argument for methodological naturalism, we should say that the argument makes use of empirical methods (the rules) to justify the legitimacy (and superiority) of those methods in philosophical research. Therefore, it is rule-circular rather than premise-circular, and viciousness of rule-circularity is far from being intuitive. Hence, the premise (VC) of the argument (AA) is false. So, the a posteriori argument for methodological naturalism is good enough to be accepted as a justification for methodological naturalism.

But, this response does not seem desirable and it has been criticized even by some naturalists. As an example, James Ladyman and Don Ross criticize preserving EDR with the aid of a similar tactic saying that even if that argument is not premise-circular and then, not viciously circular, "this style of argument will not persuade someone who totally rejects IBE." This argument can only show that using IBE in philosophy of science is 'consistent' and can be part of an 'adequate philosophy of science.'[30] Igor Douven also criticizes this tactic saying that if we allow rule-circularity, some improper reasoning rules which can be forged so that to be able to be justified using themselves will be permitted as well. For example, 'Inference to the Worst Explanation' (IWE) as an imaginary rule may be defended on the grounds that it leads to the most unsuccessful theories from the available data. Then, using IWE itself, the worst explanation of the failure of those theories is that they are true. However, it is completely irrational to conclude that IWE is a reliable rule of inference leading us to correct theories. Furthermore, Douven thinks a rule-circular argument for IBE has just the power to persuade the ones who have inclination to use IBE, not the ones who totally reject it – as Ladyman and Ross mentioned too.[31]

---

[29] Braithwaite, *Scientific Explanation*, chap. VIII; van Cleve, "Reliability, Justification," 557-559; Papineau, *Philosophical Naturalism*, 154-158.

[30] James Ladyman and Don Ross, "Scientific Realism, Constructive Empiricism, and Structuralism," in *Every Thing Must Go: Metaphysics Naturalized*, ed. James Ladyman, et al. (Oxford: Oxford University Press, 2007), 66-129, 75.

[31] See Igor Douven, "Abduction," *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta (2011), http://plato.stanford.edu/archives/spr2011/entries/abduction/ (accessed 18 August 2016), sec. 3.2.

4. Reflexivity: A Requirement for Self-Referential Theories

As I tried to show in the previous section, the naturalists' responses to the a priori argument against methodological naturalism do not seem so compelling, especially when considering the situation meta-philosophically. In this section, I put forward a proposal which, I believe, shows that the a priori argument against methodological naturalism is totally unfounded. For this reason, I focus on the requirements a philosophical theory should meet to be a candidate for solving a philosophical issue or problem in some area of philosophy or other. Generally speaking, it seems satisfactory to say that every philosophical theory should fulfill some requirements as prerequisites of being a genuine philosophical theory. A similar idea has been posed and defended about scientific theories by Larry Laudan who called such conditions 'methodological rules or standards' and tried to justify them.[32] To make it analogous with this caption, I call such a requirement for philosophical theories 'meta-methodological rule or requirement for philosophical theories' (hereafter MMR). The most general and obvious MMR – which has a parallel in science – is the requirement that every philosophical theory should be *internally consistent*. Even if some acceptable-at-face-value philosophical theories actually do not meet internal consistency requirement, there is a consensus, I think, that after their inconsistency are revealed such theories should be made consistent to be proper philosophical theories. This MMR does not seem challenging. About scientific theories, Laudan even mentions some quotations from relativists like Thomas Kuhn who endorse the consistency requirement for scientific theories.[33]

Now, I want to introduce another MMR – though it is not as general and pervasive as consistency requirement – which can be called *reflexivity requirement*. Before formulating that MMR, I should introduce three other concepts. First, *target domain* of a philosophical theory is the set of entities or n-tuple of entities which that theory is aimed to explain, reduce or unravel something about them or the relation between them – in the case of n-tuples with $n>2$.[34] As a scientific example, Darwinian theory of evolution with natural selection is about different *organisms* to explain such things as their variation, their evolutionary history and so on.[35] Therefore, its target domain is the set of all organisms in all times and places. In philosophy of mind, type identity theory is a familiar example. If we understand the

[32] Larry Laudan, *Beyond Positivism and Relativism: Theory, Method, and Evidence* (Colorado: Westview, 1996), chap. 5 and *passim*.

[33] Laudan, *Beyond Positivism*, 92.

[34] Target domain of a theory can be said to be the same thing as what Fitch calls 'subject-matter' of a theory (Frederic B. Fitch, "Self-Reference in Philosophy," *Mind* 55, 217 (1946): 64-73, 64).

[35] Fitch, "Self-Reference," 64.

theory as stating 'every mental property is identical with some physical property,' then the target domain of this theory is 2-tuples of different mental and physical properties as (M_i, P_i), and the theory is about the relation between the 2-tuples entries. And if we consider the theory as stating '*necessarily*, every mental property is identical with its related physical property,' then we should consider the 2-tuples across all possible worlds, and the target domain of the theory becomes a set like:

$\{(M_1 \text{ in } w_1, P_1 \text{ in } w_1), (M_2 \text{ in } w_1, P_2 \text{ in } w_1), \ldots, (M_1 \text{ in } w_2, P_1 \text{ in } w_2), \ldots\}$.

Second, *actual domain* of a theory is a subset of its target domain which the theory is true about its members. In practice, every theory is proposed by its theoretician as a true theory about all members of its target domain. But, it is just an ideal for every theory and not necessarily the actual case. In fact, thought experiments in philosophy are mostly aimed (as some tools) to show that a theory is not true about some (or even all) members of its target domain, or to put it another way, to show that its actual domain is not identical to its target domain. However, if actual domain of a theory is identical to its target domain, then that theory, at least at first sight, is successful.

Third, I define a *self-referential theory* as a theory that is itself directly or indirectly within its target domain (Fitch's similar definition will be introduced in the next section). I will say more about this sort of theory and its significance in philosophy in the following.

Now, I define reflexivity as a property of some philosophical theories: a theory, T, is reflexive, if and only if

> (**RC**): if T (directly or indirectly) is a member of its target domain (i.e. if it is self-referential), then it is a member of its actual domain as well.[36]

To explain (RC), I start with a trivial example. Suppose, for the sake of argument, that the claim 'all theories are false' is a philosophical theory (call it (F)). The target domain of (F) contains all theories in all areas of science, mathematics, and philosophy. Suppose (in sympathy with scientific realism) there are some true scientific theories, or instead, suppose there are at least some true mathematical theories. Then, (F) is clearly false. It is because the actual domain of (F) is not identical with its target domain. But, (F) – regarded as a theory – is itself *directly* a member of its own target domain and hence self-referential, and since it is false, it is a member of its actual domain as well. So, it is reflexive. If we change (F) into the claim 'all propositions are false' (call it (F')), then *the proposition stating* (F') will be

[36] Cf. Fitch, "Self-Reference."

a member of the target domain of (F'). So, (F') is *indirectly* self-referential.[37] According to many true propositions – from the trivial ones like 'every table is a table' to even some scientific laws – (F') and its relevant proposition are false. Therefore, it is reflexive.

Now it seems to me that reflexivity requirement, *the condition that every self-referential theory should be reflexive and fulfill (RC)*, is a genuine (though special) MMR.[38] The main reason behind the claim is that we expect every theory to embrace all the members of its target domain, and so, if that theory is a member of its own target domain (directly or indirectly), it should be true about itself as well as other members of its target domain.[39]

Reflexivity requirement so described does not seem challenging and, I think, most philosophers will endorse it. It also is not a trivial requirement. To show this, I put forward two cases in philosophy that can be regarded as evidence of upholding reflexivity requirement by some philosophers (implicitly or explicitly) as a necessary and nontrivial condition for philosophical self-referential theories.

4.1 Logical Positivists' Verification Principle: Charge of Self-Refutation

The verification principle of logical positivists was asserted by them as a criterion of meaning for sentences of any language[40], and it was posed in different forms, each of which was an endeavor to overcome the difficulties of the previous versions. In its simplest form, the principle states that a sentence is cognitively meaningful if it is analytic (or self-contradictory) – i.e. its truth or falsity can solely be judged on the meaning of its terms – or it is (at least in principle) verifiable – that is, its truth or falsity can be verified with empirical tests.[41] The principle has been criticized in different ways, one of which is intended to be discussed here: the charge of being

[37] I do not assume that a theory is combined of one or more propositions. Rather I assume, for every theory, there is a proposition (or a set of propositions) which *states* that theory. In comparison, Fitch only mentions the direct way of "self-reference." Therefore, to regard the theories concerning propositions as self-referential, he supposes, following Whitehead, that all theories are propositions and all propositions can be seen as theories (see Fitch, "Self-Reference," 65).

[38] According to (RC), if a theory is not contained in its target domain – i.e. if it is not self-referential – it is trivially reflexive, because the antecedent of (RC) is false, and so, (RC) will be true about the theory.

[39] Cf. Fitch, "Self-Reference."

[40] The criterion was initially posed for sentences. But, after a while, positivists focused on propositions rather than sentences. Here, I considered the initial version of the criterion.

[41] See, e.g., Carl G. Hempel, "The Empiricist Criterion of Meaning," in *Logical Positivism*, ed. Alfred Jules Ayer (New York: The Free Press, 1959), 108-127, 108.

self-refuting or *self-undermining*. The criticism is leveled at the principle through a question: if we apply the criterion to itself, does it itself meet the criterion, i.e. is it cognitively meaningful under itself? If it is, it should be either analytic or verifiable. If it is analytic, then it is an arbitrary principle and an empty judgment. And it cannot be verifiable regarding any concept of verifiability.[42]

Although it is stated that verificationists did not consider the challenge seriously,[43] some of the main figures of the position, including Rudolf Carnap, A. J. Ayer and Carl Hempel, met the challenge and tried to respond to it.[44] I think the criticism is relevant and the best way to explain it and the tries carried out to respond to it is via reflexivity requirement: verification principle as (part of) a theory of meaning is indirectly (via the sentence which states the principle) self-referential, and therefore, should be able to apply to itself. This fact *explains* the self-undermining charge to verification principle and the verificationists' tries to respond to it.

4.2 Strong Programme in Sociology of Scientific Knowledge

One of the main currents in Sociology of Scientific Knowledge (SSK) is the so-called Strong Programme. According to the school of thought, science should be studied by sociologists of science as a 'natural phenomenon' and as a culture or collective belief of some people (scientists, students etc.) who work within a community (the scientific community). In such studies, *whole* science as the subject matter of sociologists of science should be scrutinized rather than the only parts of science which are mostly considered false and unsuccessful; just like physiology which its aim "is to explain the organism in [both] *health and disease*."[45] Following this approach, sociologists of science do study the social, economic, and cultural conditions in which a scientific theory emerged and was accepted at a specific time in the (close or distant) history of science, and their effect on the emergence and acceptance of that theory. In brief, for proponents of Strong Programme, SSK should

---

[42] See, e.g., Hilary Putnam, *Reason, Truth and History* (Cambridge: Cambridge University Press, 1981), 106 and *passim*. Putnam tries to show that all the viewpoints like verificationism are subject to the same criticism. I think Putnam's analysis can be reconstructed with the aid of reflexivity requirement I described here.

[43] Putnam, *Reason, Truth,* 106.

[44] See Rudolf Carnap, *Philosophy and Logical Syntax* (Bristol: Thoemmes Press, 1996), 36-38; Alfred Jules Ayer, *Language, Truth, and Logic* (New York: Dover, 1952), 16; Alfred Jules Ayer, ed. *Logical Positivism* (New York: The Free Press, 1959), 15-17; Hempel, "The Empiricist Criterion of Meaning," 123-126.

[45] David Bloor, *Knowledge and Social Imagery*, 2nd ed. (Chicago: University of Chicago Press, 1991), 5 (my italics).

aim to discover the social 'causes' of scientific theories and their acceptance. Social factors are not limited to some forces from outside the scientific communities. Rather, there are many social factors within a scientific community.[46]

According to the above characterization of Strong Programme in SSK, the idea itself can be seen as a theory or approach in sociology aimed to study a 'natural phenomenon,' i.e. science. So, it is subject to this potential criticism that if we endorse it as the *correct* theory about science, then its own emergence and the expansion of its proponents must also have some social causes and social explanation. David Bloor paid attention to this criticism. But, before responding to it, Bloor had gone further introducing 'reflexivity' as one of the tenets of Strong Programme. He says about Strong Programme that: "In principle its patterns of explanation would have to be applicable to sociology itself," and he sees this requirement obligatory "because otherwise sociology would be a standing refutation of its own theories."[47] We can paraphrase Bloor's words with saying that neglecting reflexivity requirement leads Strong Programme (which is self-referential) to be excluded from its actual domain.

Bloor also takes this possibility into consideration that reflexivity requirement can be a threat to his viewpoint as an argument showing that the viewpoint is *self-refuting*. Such an argument proceeds in this way: if Strong Programme is reflexive, then part of its cause is related to social factors, and it makes Strong Programme unjustified and false. Bloor narrates some forms of this criticism from Grünwald's and Lovejoy's works.[48] Regardless of the details of Bloor's response to the charge of self-refutation, his response does not include any denial of reflexivity of Strong Programme.

5. Reflexivity vs. Consistency

There is a potential serious objection to the reflexivity requirement here.[49] The objection is that, after all, reflexivity requirement is just a special case of consistency requirement, and then, it should not be posed as an independent requirement of philosophical theories. But, I do not think it is right. I give two complementary reasons for why I think reflexivity is (and should be regarded as) an independent requirement, though the objection will not threaten my response to the a priori argument against methodological naturalism.

---

[46] See, e.g., Bloor, *Knowledge and Social*, 5-7.

[47] Bloor, *Knowledge and Social*, 7.

[48] Bloor, *Knowledge and Social*, 17-18,44.

[49] Kim Sterelny and Hossein Sheykh-Rezaee reminded me about this objection.

The first reason is that reflexivity is not the same property of theories as consistency, because a theory may be both reflexive and *inconsistent* at the same time. As an instance, the quasi-theory (F), the theory that all propositions are false, is inconsistent but reflexive.

However, it is not a plausible response to the criticism, because the reverse case is not true: a theory cannot be *irreflexive* but consistent. In fact, the main criticism is that we do not need reflexivity as an independent requirement, and consistency requirement is enough to be satisfied by theories. So, I need to give the other reason.

The second reason is not theoretical but a practical one. It seems that reflexivity is usually regarded as an independent requirement, or at least as an important special sort of consistency requirement worth mentioning as a separate requirement. If a theory is not reflexive, it is not usually said that it is inconsistent. Rather, such a theory is usually said to be *self-refuting* – as was the case about logical positivists' Verification Principle and Strong Programme. Although a self-refuting theory is inconsistent, it is inconsistent *in a special sense*; i.e. holding such a theory to be true entails its negation. A theory which is not within its target domain is not subject to self-refutation. A theory which does not have any other theory within its target domain is, a fortiori, not subject to self-refutation as well. Roughly speaking, Frederic Fitch calls the latter a 'horizontal' theory and other sort of theories 'vertical.'[50] Contrary to scientific inquiry, vertical theories are not rare in philosophy. As Fitch correctly says, in philosophical research,

> extreme comprehensiveness is sought for. Theories are constructed which purport to deal with all entities whatsoever and which therefore have an unrestrictedly extensive subject-matter. In dealing with all entities, such theories in particular deal with all theories, since theories are themselves entities of a special sort. In philosophy we thus encounter theories about the general nature of theories.[51]

He calls a vertical theory "included in its own subject-matter" – that is, in its own target domain – 'self-referential.'[52] According to him, only self-referential theories may be self-refuting, or in his words, 'self-referentially inconsistent.' He says that a horizontal theory

> may be internally inconsistent, or it may be inconsistent with known facts, and hence 'externally' inconsistent, but it cannot be inconsistent with its own nature

[50] In fact, his definitions of vertical and horizontal theories are more complicated. He introduces the concept "ordinal level" of a theory to define vertical and horizontal theories. For the exact definitions, see Fitch, "Self-Reference," 64-65.

[51] Fitch, "Self-Reference," 64-65.

[52] Fitch, "Self-Reference," 65.

> in the way that a self-referential theory can. If a self-referential theory T implies that T has the property P, and if T in fact does not have the property P, then we shall call T self-referentially inconsistent.[53]

Self-referentially inconsistent theories are, in fact, those self-referential theories which do not satisfy reflexivity requirement. So, as Fitch shows, it seems satisfactory to consider reflexivity to be a special sort of consistency. But, why should it be regarded as an independent requirement apart from consistency? The answer to this question, I think, lies in the *function* those requirements fulfill in theory appraisal.

First, to find out what the function of consistency requirement in theory appraisal is, we should understand why theories should satisfy that requirement. It seems that consistency requirement provides us with a minimal rationality: it is not rational to endorse a theory which may imply two contrary consequences. Now, I think that reflexivity requirement has a different function in theory appraisal that makes it worthy of being an independent requirement. As said above, reflexivity requirement is mainly for self-referential theories; theories like relativism,[54] positivism, some kinds of skepticism, social constructivism, any analysis of language (which essentially makes use of language itself), any theory of truth and meaning (which itself essentially bears truth value and meaning[55]), conspiracy theory, and more general and more important, any meta-philosophical doctrine about philosophy and philosophical inquiry. Such theories have usually been influential and permeating and they constitute an important part of philosophy and probably most of meta-philosophical literature. Some of these theories – like relativism and positivism – have encountered accusation of irreflexivity in practice. Sometimes, the response has been to evade reflexivity requirement with excluding the theory in question from its target domain; that, as far as I am concerned, is an *ad hoc maneuver*. Consistency requirement is too general to prevent theorists from carrying out this ad hoc maneuver in practice. There should be a more specific requirement to attract theorists' attention to avoid such a maneuver when developing their theories. Reflexivity requirement can fulfill this function. Following the requirement, a theorist will be aware that she should develop her self-referential theory so that it includes itself in its actual domain, or her theory should not be self-referential at all, albeit she should not appeal to an ad hoc maneuver to exclude her theory from its target domain.

---

[53] Fitch, "Self-Reference," 66. The property P may be said to be the property that T ascribes to all the theories in its target domain.

[54] See Maria Baghramian, *Relativism* (London: Routledge, 2004), 100-107.

[55] See, e.g., Roth, "Theories of Nature."

6. Methodological Naturalism and Reflexivity Requirement

Now, if we see methodological naturalism as a philosophical theory expressed through (MN), what is its target domain? To answer the question, it should be made clear what (MN) is about. (MN) prescribes following scientific method for philosophers. But, what should we do if we are to assess an *existing* piece of work or a theory in philosophy to see whether and to what extent it is naturalistic in this sense? To put it another way, what works in philosophy do naturalists endorse as naturalistic? If naturalists are asked to name some naturalistic works in philosophy, they will definitely cite, inter alia, evolutionary works (in different areas of philosophy like epistemology, ethics and philosophy of science) as their paradigm cases of naturalistic works in philosophy.[56] Now, what does it make such theories naturalistic? It does not seem enough for a thoroughgoing naturalistic work to just be inspired or informed by a scientific theory. For if, say, evolutionary epistemologists justified their view *traditionally* using armchair methods, their view would be no longer considered to be naturalistic. What makes, say, a theory in evolutionary epistemology of theories (EET) naturalistic is, I think, the assertion that it is (as claimed) *confirmed* with relevant evidence of the *actual* changes of theories in the past history of science and it will probably be confirmed in the future as well – though under a general sense of 'confirmation.' It means that EET to be naturalistic should be justified abductively, using IBE as a scientific method.

If I am right, it seems that (MN) can be rewritten as follows:

> (**MN′**): Any philosophical theory can constitute knowledge only if it is justified scientifically (i.e. through scientific method).[57]

According to (MN′), the target domain of methodological naturalism is the set of all philosophical theories within which, there is methodological naturalism itself, and so, it is self-referential.[58] Hence, methodological naturalism to be reflexive should be true about itself, that is, it should be justified through scientific method. Therefore, not only does an a posteriori argument for methodological naturalism not lead to a vicious circle, but also it shows that methodological naturalism fulfills reflexivity requirement through such an argument.

There are some clues in the works of anti-naturalists about this matter. In fact, it seems that some anti-naturalists take reflexivity requirement (implicitly) for

---

[56] See Rosenberg, "A Field Guide."

[57] See Elliott Sober, *Ockham's Razors* (Cambridge: Cambridge University Press, 2015), 245.

[58] Fitch makes a similar point when he says: "consider the view that every valid theory must be obtained from observed empirical data. This is a theory about theories and their validity" (Fitch, "Self-Reference," 66).

granted and try to show that methodological naturalism is not reflexive since it can only have an a priori justification if it is justified at all. The criticism can be traced back to Sheldon's 1945 article. The quotation below is one of his attacks to '1944 school of naturalism:'[59]

> Your own creed tells you not to believe anything till it is experimentally confirmed. How can you expect us to believe you have the right method for philosophy until you show us that it succeeds in giving us objective truth comparable with that of the sciences, truth on which the philosophic experts agree? […] I fear that you claim to have proved the rightness of your method in philosophy beforehand; an a priori claim you should be the last to make.[60]

As is immediately apparent in the last sentence of the above quotation, Sheldon states that naturalists to be coherent (better to say, reflexive) are not allowed to appeal to a priori defense of naturalism. This is the first horn of naturalist's dilemma.

Williamson also drops a similar hint saying that this claim of radical naturalists that "all truths are discoverable by hard science" is expected to be discoverable through hard science. But it is not, Williamson claims, and so, this kind of naturalism is not true.[61]

Fitch also touches on the point. About "the view that every valid theory must be obtained from observed empirical data," he writes:

> Incidentally it is a theory which does not seem to conform to its own criterion as to what constitutes a valid theory, at least not unless it can itself be shown to have been obtained as a generalization from observed empirical data.[62]

At the end of this section, another point should be added. The methodological traditionalism is also reflexive because it is justified through traditional methods. So, the main point remains here is that how we can compare these two methodologies in philosophy; the issue that will be addressed in the next section.

7. Reflexivity and Justification

What I have tried to show up to this point is that having an a posteriori argument for methodological naturalism does not make it viciously circular at all, and rather, according to reflexivity requirement for self-referential theories, such an argument for naturalism is mandatory and satisfying the requirement should be considered a

[59] See Kim, "American Origins," 86; Sheldon, "Critique," 253.
[60] Sheldon, "Critique," 268.
[61] Williamson, "Unclarity," 37. See also BonJour, "Against," 7.
[62] Fitch, "Self-Reference," 66. See also Siegel, "Naturalized Epistemology," 57-58; Siegel, "Empirical Psychology," 675-676.

merit for methodological naturalism, and then, traditionalist a priori argument against methodological naturalism is totally unfounded.

But now, there remains a question here: does reflexivity of a theory confer justification on it? I do not suppose it does. The main reason behind this is the fact that being reflexive means that a theory is true about one (though a special) member of its target domain, not all (or even most) of the members of its target domain comprehensively. Therefore, being reflexive does not confer an *adequate* justification on any theory and a theory should be justified independently of being reflexive.

But the problem with methodological naturalism is specifically thornier. It is because of the fact that any kind of justification proceeds following a special method. And methodological naturalism is a meta-philosophical idea (or norm) about the legitimate methods of philosophical inquiry. Now, which method should be adopted to justify a methodological view like methodological naturalism? Papineau, tackling with the charge of circularity, drops a hint about this matter: such an accusation can be leveled against traditionalism since it is a methodological view as well.[63] In fact, we have two competing views each of which has its own methodology, and as a result, each prescribes a kind of justification according to its own method. So, the dispute apparently ends in stalemate and we are left with no choice but to admit those views are *incommensurable in a sense*, i.e. methodologically incommensurable.

Incommensurability of scientific theories has been a problem after the works of Thomas Kunh and Paul Feyerabend who infamously known as the founders of modern relativism. But there are some solutions to this problem, posed by philosophers of science like Imre Lakatos and Laudan. Now, if the problem about scientific theories has some (for me, compelling) solution, why not try it on the parallel problem about philosophical theories? I believe that the best solution to the problem of incommensurability about scientific theories (which I think is Laudan's) can in the same way be employed to address the problem of incommensurability of philosophical methodologies. There is not enough space here to elaborate the view, but I introduce the main elements of the suggestion.

Laudan in addressing issues like scientific progress and incommensurability of scientific theories in history of science, proposed his concept of *research tradition* – similar to but not identical with Kuhn's concept of 'paradigm.' A research tradition is a more general view than a single theory, that includes ontological and methodological parts and overarches theories which follow the ontology and methodology of the tradition. Therefore, two theories from two different research

[63] See Papineau, *Philosophical Naturalism*, 157-158.

traditions even differ on their methods. Such theories cannot (and should not) be compared with each other directly. Instead, we can compare their traditions (even if they are incommensurable considering their basic ontological entities and/or their research methods). Competing traditions are compared according to their *problem-solving effectiveness*, that is, the total number of problems solved by (the theories of) each of them (with weighing their importance).[64] Problem-solving effectiveness of a tradition shows the success (or failure) of it: the tradition with more amount of problem-solving effectiveness – that which solves more problems (and the most important ones) – is more successful. Progress rate of a tradition can be measured by the problem-solving rate of that tradition, and we should also invest in traditions which show high progress rates awhile after their appearance.[65]

This *model* works in philosophy as well. Such views like naturalism and traditionalism, I think, are not philosophical theories or theses; rather, they are research traditions. Naturalism, for example, has its own methodology and ontological constraints (the latter due to *ontological* naturalism), and many philosophical theories follow its methodology and ontology. Therefore, naturalism and traditionalism should be assessed and compared to each other according to their success in problem-solving. If this model is pertinent, the only way for justifying naturalism (and traditionalism as well) is appraising its success in problem-solving in comparison with its rival, rather than trying to adduce a philosophical argument for (or against) it. I hope to articulate this view in another occasion.[66]

---

[64] In measuring the problem-solving effectiveness of a research tradition, we should also take into account the *conceptual problems* the tradition poses. The number of conceptual problems (considering their importance) decreases the problem-solving effectiveness of a tradition (Larry Laudan, *Progress and Its Problems: Towards a Theory of Scientific Growth* (Berkeley, CA: University of California Press, 1977), chaps. 2-3).

[65] For more details, see Laudan, *Progress*. For a brief introduction to his proposal, see Laudan, *Beyond Positivism*, chap. 4.

[66] The research project that led to this article was funded by Iranian Institute of Philosophy. Part of the research has been carried out at the Australian National University. I should thank the officials of ANU School of Philosophy and Research School of Social Sciences for letting me stay at the quiet office which enabled me to conduct the research. I am deeply indebted to Kim Sterelny, Hossein Sheykh-Rezaee, and Laleh Ghadakpour, who lend me a great deal of support in my research project. I should also thank Alexander Rosenberg, Alexander Sandgren, and Stephen Mann for their valuable comments on earlier drafts of this paper.

CONSISTENCY AND SHIFTS IN GETTIER CASES

Andreas STEPHENS

ABSTRACT: Two Gettier cases are described in detail and it is shown how they unfold in terms of reflective and reflexive desiderata. It is argued that the Gettier problem does not pose a problem for conceptions of knowledge as long as we are consistent in how we understand justification and knowledge. It is only by reading the cases with a reflective understanding of justification but a reflexive understanding of knowledge, without acknowledging that this takes place, that the cases become 'problems.'

KEYWORDS: Gettier, knowledge, justification, intuition, reflective processes, reflexive processes, consistency

1. Introduction

The paradigmatic definition of declarative propositional knowledge states that justification, truth, and belief (JTB) are individually necessary and jointly sufficient conditions for knowledge. But, in his widely influential article 'Is Justified True Belief Knowledge?'[1] Edmund Gettier questions this definition and instead argues that JTB is insufficient for knowledge:

> (G): JTB is not sufficient as a definition of knowledge.

Gettier presents two premises for his analysis:

> (A): It is possible for *S* to be justified in his belief that *p* based on a false proposition.
>
> (B): If *S* is justified in his belief that *p*, and *q* follows from *p*, and *S* deduces *q* from *p* and accepts *q* as a result of the deduction then *S* is justified in his belief that *q*.

Two counterexamples against JTB are then offered in support of (G) – and many others have emerged – where a subject is presented as being justified in his true belief that *x* but we, supposedly, are unwilling to accept that the subject *knows* that *x*.

To exemplify, Gettier's first counterexample (Case I) presents a situation where Smith and Jones have applied for a job. In a first step, we are told that the president of the company has told Smith that Jones will get the job, and Smith has

[1] Edmund L. Gettier, "Is Justified True Belief Knowledge?," *Analysis* 23, 6 (1963): 121-123.

recently counted the coins in Jones' pocket. Smith is hence, supposedly, justified in his belief that:

(1): Jones will get the job and Jones has ten coins in his pocket.

Based on (1), Smith deduces (2) and is thus, supposedly, justified in his belief that:

(2): The person who will get the job has ten coins in his pocket.

But as it turns out, in a second step, it is revealed that Smith gets the job. And finally, in a third step due to a coincidence, it is revealed that he has ten coins in his pocket. So, (2) is true even though (1) is false. Smith believes that (2) is true and Smith is, supposedly, justified in his belief that (2) is true. All parts of JTB are hence satisfied, but we are, according to Gettier, unwilling to declare that Smith knows that (2), which would support (G).

There is a wide taxonomy of different ways to formulate Gettier cases, as pointed out by for example Blouw, Buckwalter, and Turri,[2] but I will, below, focus on the aforementioned influential case (I) by Gettier and an additional (fake barn) case by Goldman.[3]

In section 2, I discuss where intuition-based approaches, as well as experimental philosophical approaches, seemingly leave us. In section 3, I then stepwise explore the formulation and case set-up of two paradigmatic Gettier cases focusing on a reflective (subject-centered, internalist) reading and a reflexive (evaluator-centered, externalist) reading to elucidate 'the problem.'

2. A Smorgasbord of Intuitions

Ever since the article appeared, epistemologists have tried to tackle Gettier's argument and his counterexamples. For example, Dretske[4] describes how someone who has 'conclusive reasons' cannot have grounded her belief on false evidence. Lehrer and Paxon[5] reinterpret the traditional definition JTB and add an extra clause of undefeatability. Another version of this undefeatability-form of amendment is

[2] Peter Blouw, Wesley Buckwalter, and John Turri, "Gettier Cases: A Taxonomy," in *Explaining Knowledge: New Essays on the Gettier Problem*, eds. Rodrigo Borges, Claudio de Almeida, and Peter D. Klein (Oxford: Oxford University Press, 2017): 242-252.

[3] Alvin Goldman, "Discrimination and Perceptual Knowledge," *The Journal of Philosophy* 73, 20 (1976): 771-791.

[4] Frederick Dretske, "Conclusive Reasons," *The Australasian Journal of Philosophy* 49 (1971): 1-22.

[5] Keith Lehrer and Thomas Paxon Jr., "Knowledge: Undefeated Justified True Belief," *The Journal of Philosophy* 66 (1969): 225-237.

Goldman's[6] 'no-relevant-alternatives' condition, which focuses on the causality and reliability of the processes under consideration. Using a logical approach, Floridi[7] claims that Gettier does not succeed with his counterexamples since JTB is in principle irreparable, and similarly Zagzebski[8] claims that both approaches that amend JTB with a fourth condition and approaches that strengthen the justification condition necessarily fail. Williamson[9] has even suggested that attempts of analyzing knowledge should be discontinued and that a 'knowledge-first' approach should be adopted instead.

Importantly, all the aforementioned lines of inquiry – including Gettier's article – are positioned in what can broadly be construed as the conceptual analytic philosophical tradition, focusing on language and intuitions. However, there have also been a number of attempts to get to the bottom of the issue by empirically exploring just what peoples' intuitions amount to. Thus, several experiments have been conducted where early experimental philosophical findings suggested that intuitions about when knowledge and justification obtains, related to a number of epistemological cases including Gettier cases, systematically differ between various groups.[10] Focusing on Gettier's problem, a number of cultural specific differences were found. Furthermore, intuitions between different socioeconomic groups were shown to vary significantly.[11] But, various later studies have claimed that these findings do not withstand closer scrutiny.[12] Moreover, it has been argued that

[6] Goldman, "Discrimination;" Alvin Goldman, *Epistemology and Cognition* (Cambridge: Harvard University Press, 1986).

[7] Luciano Floridi, *The Philosophy of Information* (Oxford: Oxford University Press, 2011).

[8] Linda Zagzebski, "The Inescapability of Gettier Problems," *The Philosophical Quarterly* 44, 174 (1994): 65-73.

[9] Timothy Williamson, *Knowledge and its Limits* (Oxford: Oxford University Press, 2000).

[10] Jonathan Weinberg, Chad Gonnerman, Cameron Buckner, and Joshua Alexander, "Are Philosophers Expert Intuiters?," *Philosophical Psychology* 23, 3 (2010): 331-355; Jonathan Weinberg, Shaun Nichols, and Stephen Stich, "Normativity and Epistemic Intuitions," *Philosophical Topics* 29, 1-2 (2001): 429-460; Joshua Knobe and Shaun Nichols, eds., *Experimental Philosophy* (Oxford: Oxford University Press, 2008); Joshua Alexander, *Experimental Philosophy: An Introduction* (Cambridge: Polity, 2012).

[11] Weinberg, Nichols, and Stich, "Normativity and Epistemic Intuitions."

[12] Minsun Kim and Yuan Yuan, "No Cross-Cultural Differences in the Gettier Car Case Intuition: A Replication Study of Weinberg et al. 2001," *Episteme* 12, 3 (2015): 355-361; Hamid Seyedsayamdost, "On Gender and Philosophical Intuition: Failure of Replication and Other Negative Results," *Philosophical Psychology* 28, 5 (2015): 642-673; Hamid Seyedsayamdost, "On Normativity and Epistemic Intuitions: Failure of Replication," *Episteme* 12, 1 (2015): 95-116; John Turri, "Knowledge Judgments in 'Gettier' Cases," in *A Companion to Experimental Philosophy*, eds. Justin Sytsma and Wesley Buckwalter (Malden, Mass.: Wiley-Blackwell, 2016): 337-348.

epistemic intuitions might be more reliable and similar than previously supposed.[13] Turri[14] even claims that laypeople and professional philosophers of different cultural backgrounds, ages, and genders to a large extent share the intuition of accepting (G). To reach this conclusion Turri has conducted experiments that stage-wise 'guides' subjects through the cases under investigation. This guidance, according to Turri, allows 'laypeople to competently assess Gettier cases,'[15] although such guidance, in my view, is problematic since it might be hard to avoid that test-subjects are influenced to reach a conclusion that suits the experimenter.

So, it remains an open question whether intuitions differ in *systematic* ways, but the sheer amount, and variation, of interpretations about the Gettier problem that has been found in different empirical experiments – as well as those that can be found in the vast Gettier corpus – strongly indicates, or so I argue, that experimental philosophy have, so far, failed to generate closure in this debate that still is alive after more than fifty years.

For discussions about how the fruitfulness of JTB might trump the potential difficulties Gettier introduces see Weatherson,[16] Lozanski,[17] and Olsson.[18] For a discussion about the overarching value of the Gettier problem and the philosophical discussions concerning it see Turri.[19] There are also discussions pertaining to, for example, the possibility of better/correct and worse/incorrect intuitions.[20] But it remains dubious whether it is possible to find reasonable grounds to motivate any specific choice regarding who's intuitions ought to be heard or ignored.

In short, no universally endorsed position can be found regarding what the lesson from Gettier is, how the cases should be interpreted, if the problem should be

---

[13] Kenneth Boyd and Jeniffer Nagel, "The Reliability of Epistemic Intuitions," in *Current Controversies in Experimental Philosophy,* eds. Edouard Machery and Elizabeth O'Neill (New York: Routledge, 2014): 109-127; Edouard Machery, Stephen Stich, David Rose, Amita Chatterjee, Kaori Karasawa, Noel Struchiner, Smita Sirker, Naoki Usui, and Takaaki Hashimoto, "Gettier Across Cultures," *Noûs* 51, 3 (2017): 645-664.

[14] John Turri, "A Conspicuous Art: Putting Gettier to the Test," *Philosopher's Imprint* 13, 10 (2013): 1-37.

[15] Turri, "A Conspicuous Art," 34.

[16] Brian Weatherson, "What Good are Counterexamples?," *Philosophical Studies* 115, 1 (2003): 1-31.

[17] Lukasz Lozanski, "The Gettier Problem," *Philosophy Now* 63 (2007): 28-29.

[18] Erik J. Olsson, "Gettier and the Method of Explication: A 60 Year Old Solution to a 50 Year Old Problem," *Philosophical Studies* 1, 172 (2015): 57-72.

[19] John Turri, "Manifest Failure: The Gettier Problem Solved," *Philosopher's Imprint* 11, 8 (2011): 1-11; Turri, "Knowledge Judgments."

[20] Alexander, *Experimental Philosophy*; Weinberg, Gonnerman, Buckner, and Alexander, "Are Philosophers Expert."

solved or dissolved, or if – and if so how – JTB can be salvaged. There seemingly are as many theories and intuitions around as there are theorists.

3. Two Readings of Gettier Cases

As aforementioned, Gettier cases have frequently been viewed as being problematic, forcing numerous theoreticians into reformulating their views and definitions of knowledge. Inspired by Kaplan's[21] discussion, I will in this section explore how two well-known Gettier cases come across from a reflective (subject-centered, internalist) and a reflexive (evaluator-centered, externalist) perspective on justification and knowledge.[22] This is accomplished by delineating the Gettier problem with a focus on the *consistency* of one's reading. In order to retain readability, I have limited my exploration to two illustrative Gettier cases although I believe that similar elucidations can be given concerning other formulations and forms of Gettier cases.

3.1 Careers and Coins

I here present Gettier's[23] first case in full:

> *Case I:*
>
> Suppose that Smith and Jones have applied for a certain job. And suppose that Smith has strong evidence for the following conjunctive proposition:
>
> (1) Jones is the man who will get the job, and Jones has ten coins in his pocket.
>
> Smith's evidence for (1) might be that the president of the company assured him that Jones would in the end be selected, and that he, Smith, had counted the coins in Jones's pocket ten minutes ago. Proposition (1) entails:
>
> (2) The man who will get the job has ten coins in his pocket.
>
> Let us suppose that Smith sees the entailment from (1) to (2), and accepts (2) on the grounds of (1), for which he has strong evidence. In this case, Smith is clearly justified in believing that (2) is true.
>
> But imagine, further, that unknown to Smith, he himself, not Jones, will get the job. And, also, unknown to Smith, he himself has ten coins in his pocket.

[21] Mark Kaplan, "It Is Not What You Know That Counts," *The Journal of Philosophy* 82, 7 (1985): 350-363.

[22] George Pappas, "Internalist vs. Externalist Conceptions of Epistemic Justification," in *The Stanford Encyclopedia of Philosophy* (Fall 2017 Edition), ed. Edward N. Zalta (Stanford: Stanford, CA, 2017): URL = https://plato.stanford.edu/archives/fall2017/entries/justep-intext/.

[23] In what follows, I will repeatedly return to this quote from Gettier, "Is Justified," 122. For clarity of presentation, I have changed Gettier's two claims '(d)' and '(e)' into '(1)' and '(2).'

> Proposition (2) is then true, though proposition (1), from which Smith inferred (2), is false. In our example, then, all of the following are true: (*i*) (2) is true, (*ii*) Smith believes that (2) is true, and (*iii*) Smith is justified in believing that (2) is true. But it is equally clear that Smith does not *know* that (2) is true; for (2) is true in virtue of the number of coins in Smith's pocket, while Smith does not know how many coins are in Smith's pocket, and bases his belief in (2) on a count of the coins in Jones's pocket, whom he falsely believes to be the man who will get the job.

3.1.1 First step

Gettier's first counterexample initially outlines a rather mundane situation:

> Suppose that Smith and Jones have applied for a certain job. And suppose that Smith has strong evidence for the following conjunctive proposition:
>
> (1) Jones is the man who will get the job, and Jones has ten coins in his pocket.
>
> Smith's evidence for (1) might be that the president of the company assured him that Jones would in the end be selected, and that he, Smith, had counted the coins in Jones's pocket ten minutes ago. Proposition (1) entails:
>
> (2) The man who will get the job has ten coins in his pocket.
>
> Let us suppose that Smith sees the entailment from (1) to (2), and accepts (2) on the grounds of (1), for which he has strong evidence. In this case, Smith is clearly justified in believing that (2) is true.

The president of the company has had a casual talk with Smith telling him that Jones will get the job, and for some reason Smith has counted Jones' pocket-change. From this Smith concludes that (1) – Jones will get the job and Jones has ten coins in his pocket – and then extrapolates and accepts (2).

A reflective point of view will make an evaluator focus on Smith's reflective processes such as attention, information-manipulation, reasoning, and decision making, whether he can be seen as epistemically responsible and rational in holding his belief, and whether he has cognitive access to his belief. Since the counterexample, initially, describes a commonplace situation where the president of the company has talked to Smith and told him that Jones will get the job, and Smith has counted Jones' coins, an interpretation that holds Smith epistemically responsible and justified in holding his belief that (1) is true is indeed plausible. Given the evidence Smith has, as Gettier presents the case, Smith is acting rationally in holding his belief. And since the characteristic point of view, from a reflective perspective, is the first-person point of view of Smith, no more, could reasonably be demanded from a person in Smith's position. Smith has cognitive access to his belief and can through reflection, reasoning, and information-manipulation form arguments in favor of his conclusion that (1) is true, on which he then can base

decisions. The evidence or justification Smith has, viewed reflectively, is nothing out of the ordinary but is nevertheless described as 'strong' which could be interpreted as indicating that the justification needed for knowledge can be met or fulfilled to a high enough degree.

Given what an evaluator knows of the situation in the counterexample up until this point, and given that her focus is a reflective reading, thus makes it plausible of her to interpret the situation in a way that we rightfully can say that Smith knows that (2) is true. The initial step of the argument hence presents a situation where Smith is justified (by his reflective processes) and where JTB is sufficient as a definition of knowledge.

If an evaluator instead views the initial step of the case from a reflexive point of view that centers on reflexive processes such as pattern recognition the counterexample makes her focus on whether Smith has gotten his belief in a reliable way, through a reliable process, and whether he is favorably (causally) connected to the world. It is thus the reflexive grounds that Smith has for his justification and knowledge that is relevant. As the counterexample initially is laid out, Smith's conversation with the president of the company and his investigation of Jones' pockets seems to be reliable processes. Smith has taken part of an everyday interaction in which he is likely to correctly pick up on relevant patterns. The characteristic point of view, from a reflexive perspective, is the third-person, and no more seems reasonable of an evaluator to demand from the described situation. Given the everyday formulation of the counterexample and since Smith's evidence is described as 'strong' it is plausible to interpret the situation in a way that makes Smith's conclusion that (1) true. Smith has used a reliable process to form his belief. Everything, so far, points towards that Smith is favorably connected to the world.

Given what an evaluator knows of the situation in the counterexample up until this point, and given that her focus is reflexive, thus makes it plausible of her to interpret the situation in a way that she rightfully can say that Smith knows that (2) is true. The initial step of the argument hence presents a situation where Smith is justified and where JTB is sufficient as a definition of knowledge.

It should however be pointed out that the initial step of the counterexample only describes the way and processes Smith has used to justify his belief. These processes seem sufficient, if charitably interpreted, since they are described as giving rise to strong evidence and since the situation seems ordinary. But, notably, Gettier does not, at this point, explicitly describe Smith's causal connection to the world at this step. Rather an evaluator has to choose how to interpret the presented situation.

3.1.2 Second Step

In the following step of his counterexample, after the deduction to (2), Gettier presents a new development:

> But imagine, further, that unknown to Smith, he himself, not Jones, will get the job.

It is Smith that gets the job. But since Smith has not yet been informed of this, it is, from a reflective evaluative perspective, still reasonable for him to believe that it is Jones who will get the job. If the sole focus is on Smith's first-person point of view and reflective processes of active conscious reasoning – he therefore still knows that (2) is true. But since the new situation bluntly states that (2) is false, this conclusion of course seems problematic from an objective third-person perspective. Nevertheless, if an evaluator's only focus is on reflective processes, Smith knows.

Seen from a reflexive perspective, the described situation in the second step puts Smith's processes' reliability in question. If an evaluator demands absolute certainty, they should be considered unreliable. But, if she allows fallibilism, which perhaps is a more plausible interpretation, it is fine that reliable processes sometimes – especially under less common situations – come out false. Thus, there is a vagueness regarding how the situation should be interpreted.

However, as is made clear if the evaluator focuses on the truth-connect, there is an outright falseness involved. As aforementioned, Gettier left this aspect open in his initial formulation. He now fills in some details and the world does not support Smith's initial beliefs. It is not Jones who will get the job. But once again, Gettier leaves an aspect unspecified. On the one hand, an evaluator has been explicitly told that it is Smith – not Jones – who will get the job. But nothing is mentioned regarding coins, and since it might seem to be an irrelevant technicality to the situation as a whole it might be easy to overlook and ignore this detail.

Depending on how an evaluator chose to see the situation, Smith will either be justified or not, and accordingly either know that (2) is true or not.

3.1.3 Third Step

In the third step, Gettier describes how it not only is Smith that gets the job, but, due to a coincidence, it turns out that he has ten coins in his pocket:

> And, also, unknown to Smith, he himself has ten coins in his pocket.

The evidence Smith has, no longer seems strong but instead irrelevant and coincidental in relation to (2). But even given this new information, Smith still seems to be reflectively justified, since nothing has changed from *his* perspective. It is

therefore still reasonable and rational of him that he should believe what he does from a point of view that focuses on reflective processes. Taking the whole counterexample, and the situation such as Smith sees it into account, Smith is still justified, from his first-person point of view, and thus in a situation where JTB is satisfied and he therefore knows that (2) is true.

From a reflexive perspective, the counterexample's third step is, just as the second step, formulated in a way that makes it ambiguous whether Smith is justified. He can be seen as being causally connected to the world in a favorable way; the person who will get the job, indeed, has ten coins in his pocket. So with this in mind, JTB is satisfied and Smith does know that (2) is true. However, the reflexive processes Smith has used can be considered unreliable since they seem to only involve irrelevant and coincidental information with respect to the situation. On such an interpretation, JTB is not satisfied and Smith does not know that (2) is true. But it is also possible to interpret the described situation to involve reliable processes that only happen to come out wrong given the extraordinary situation. Then JTB is satisfied and Smith does know that (2) is true. Depending on whether an evaluator focuses on the reliability of Smith's process or his truth-connect, different interpretations are thus possible.

3.1.4 Careers and Coins: Discussion

As has been shown, the counterexample can be interpreted in more than one way, and it is thus possible to reach different conclusions regarding whether Smith knows that (2) is true. Smith initially seems to be in a situation where JTB is satisfied and he knows whether the focus is on reflective or reflexive processes. But in the second and third steps of the argument, interpretations regarding this matter can come apart depending on which process-form, and interpretation, an evaluator chooses to focus on. However, (G) does not follow for consistent readings. Either Smith is justified and knows, or he is not justified and does not know.

Gettier continues his article with a second counterexample. However, it is in all essentials identical to the first. An analysis of this counterexample hence leads, *mutatis mutandis*, to the same result as the first counterexample. Rather than repeating myself I will therefore instead investigate an additional, and importantly dissimilar, counterexample offered by Alvin I. Goldman.

3.2 Real and Fake Barns

Goldman[24] presents a case that differs importantly from the previous counterexample regarding the reflexive perspective, and I will thus only address that aspect. I will treat Goldman's case as a Gettier case even though it is part of a discussion concerning a more general theory of 'knowing' pursued by Goldman. In defense of this usage it is worth pointing out that Goldman, in direct connection to the presented situation, does state that 'the traditional justified-true-belief account of knowledge is of no help in explaining this change.'[25]

3.2.1 First Step

A perfectly ordinary situation is described:

> Henry is driving in the countryside [… and] identifies various objects on the landscape as they come into view. […] Henry has no doubt about the identity of these objects; in particular, he has no doubt that the last-mentioned object is a barn, which indeed it is. Each of the identified objects has features characteristic of its type. Moreover, each object is fully in view, Henry has excellent eyesight, and he has enough time to look at them reasonably carefully, since there is little traffic to distract him. Given this information, would we say that Henry *knows* that the object is a barn? Most of us would have little hesitation in saying this, so long as we were not in a certain philosophical frame of mind.[26]

3.2.2 Second Step

Once again a shift is introduced:

> Suppose we are told that, unknown to Henry, the district he has just entered is full of papier-mâché facsimiles of barns. These facsimiles look from the road exactly like barns, but are really just façades, without back walls or interiors, quite incapable of being used as barns. They are so cleverly constructed that travellers invariably mistake them for barns. Having just entered the district, Henry has not encountered any facsimiles; the object he sees is a genuine barn. But if the object on that site were a facsimile, Henry would mistake it for a barn. Given this new information, we would be strongly inclined to withdraw the claim that Henry *knows* the object is a barn. How is this change in our assessment to be explained?[27]

---

[24] Goldman, "Discrimination."

[25] Goldman, "Discrimination," 773; Goldman formulates his view on knowledge in the following words: '[A] person is said to know that *p* just in case he *distinguishes* or *discriminates* the truth of *p* from relevant alternatives' (Alvin Goldman, "Discrimination," 772).

[26] Goldman, "Discrimination," 772.

[27] Goldman, "Discrimination," 773.

In the second step of Goldman's argument, the occurrence of facsimiles is mentioned – which is similar to the way the aforementioned case is laid out. But a change compared to the previous counterexample should be noted. Henry is causally connected to the world in a favorable way – it is an actual barn that he sees. But, whereas it, in the previous case, was an irrelevant coincidence that saw to the world-connection, it is in this case unproblematic. It is instead the surrounding context that might instill a sense of ambiguity in an evaluator since it consists of possible 'defeaters.' Nevertheless, since there is a truth-connect between Henry and the world, the reflexive processes he uses are reliable and JTB should be seen as being satisfied.

3.2.3 Real and Fake Barns: Discussion

According to Goldman the new situation presented in the second step would make an evaluator inclined to withdraw the claim that Henry knows that the object is a barn, and he asks how this change in assessment is to be explained. It could be debated whether this is necessarily counterintuitive or not. But if we accept that there is a problem here – how can it be elucidated?

One way of making sense of this view is to point out that the counterexample's shift questions the reliability of the reflexive process Henry has used to attain his evidence. In an area filled with facsimiles it might no longer be enough to just look at objects from a distance, it might instead be necessary to investigate further. The traffic in the area might no longer be ignored since it perhaps will require absolute focus to reliably separate a real barn from a fake barn. The processes Henry uses might no longer plausibly be interpreted as reliable, because of the accidental new feature attached to them. But even so, Henry is in fact favorably causally connected to the world, and it is rather the context that involves an element of unreliability than the specific process he used. This formulation might thus give rise to conflicting interpretations if evaluators are unclear as to which aspect of the reflexive processes they prioritize as being most important.

The described situation is however not mysterious in any deeper sense. Under ecologically normal situations the reflexive processes Henry uses, at least on a charitable reading, are plausibly seen as being reliable. In the described contrived situation, this is no longer the case, and so Henry would not know that what he sees is a barn. But, as long as Henry doesn't spend too much time in this new outlandish context he could reasonably still be considered to use a reliable process – albeit a fallible one – and thus know that what he sees is a barn.

3.3 The Problem-creating Shift?

Focusing on Gettier's first counterexample, a consistent reflective reading gives that Smith is reflectively justified and knows that (2) is true. (G) would then not follow. Concerning a reflexive reading, the situation is slightly more complicated. On the one hand there is a truth-connect, and if seen as representative of a typical situation the processes involved seem reliable. Accordingly, from a consistent reading that heeds these commitments Smith is justified and knows that (2) is true, and so (G) would not follow. On the other hand, the relevant reflexive processes' reliability is questionable. A consistent reading from this outlook would then conclude that Smith is neither justified nor knows that (2) is true, and so (G) would not follow.

Depending on whether an evaluator focuses on reflective processes or reflexive processes – some form of reflexive reading – she will thus reach different conclusions regarding whether Smith is justified or knows that (2) is true. However, (G) does not follow for consistent readings, and so, Gettier cases do not undermine consistent analyses of justification and knowledge.

But, importantly, Gettier's case is instead formulated in a way that, arguably, makes most sense if its initial step is read from a reflective perspective. Especially premise (A) fits rather uneasily with a reflexive perspective, as well as seeing Smith as being 'clearly justified' based on the presented evidence. The second and third steps instead make more sense from a reflexive perspective – where the lack of a truth-connect is highlighted concerning (1), followed by a newly introduced situation of relevance for (2). But, when Gettier then writes the following:

> In our example, then, all of the following are true: (*i*) (2) is true, (*ii*) Smith believes that (2) is true, and (*iii*) Smith is justified in believing that (2) is true.

He seems to have gone back to a reflective understanding of justification. However, his understanding of knowledge is, arguably, reflexive since he concludes:

> But it is equally clear that Smith does not know that (2) is true; for (2) is true in virtue of the number of coins in Smith's pocket, while Smith does not know how many coins are in Smith's pocket, and bases his belief in (2) on a count of the coins in Jones's pocket, whom he falsely believes to be the man who will get the job.

It thus seems that in order to accept all steps of Gettier's argument – which leads to (G) – the case must be read with a reflective understanding of justification but a reflexive understanding of knowledge.

4. Concluding Remarks

It has been argued that as long as evaluators of Gettier cases are clear about which positions they choose to take concerning justification and knowledge – if their reading is reflective or reflexive – different conclusions will be reached. However, on all consistent readings, (G) does not follow and so Gettier cases do not need to undermine analyses of justification and knowledge, or be seen as genuine problems. It is by allowing the vagueness, and shifting focus back and forth, in the cases to influence one's evaluation into an inconsistent reading that (G) becomes an option.

This said, it is not clear that JTB is a fruitful definition of declarative propositional knowledge. It can also be argued that the above discussion indicates that a naturalist approach to investigating knowledge is preferable, in that its methodology is not plagued by the vagueness of intuition-based inquiries. Furthermore, it could be claimed that justification indeed should be seen as involving something reflective whereas knowledge involves something reflexive, although in that case this ought to be clearly stated to avoid confusion, and would render Gettier cases unproblematic – as long as one accepts this view.

Lastly, I want to highlight that even though it probably is possible to compose new Gettier cases *ad infinitum*, a lot seems to have been gained if an evaluator, when faced with each new problem, can ask herself how this problem handles reflective processes and reflexive processes.[28]

[28] Acknowledgements: I want to thank Martin L. Jönsson who has been exceedingly generous with his knowledge, eye for detail, and time. Thanks to Erik J. Olsson, Ylwa Sjölin Wirling, Asger Kirkeby-Hinrup, participants at the higher seminar in theoretical philosophy at Lund University, and my anonymous reviewers for insightful comments.

WHAT IS THE RELATION BETWEEN SEMANTIC AND SUBSTANTIVE EPISTEMIC CONTEXTUALISM?

Ron WILBURN

ABSTRACT: Epistemic Contextualism is generally treated as a semantic thesis that may or may not have epistemological consequences. It is sometimes taken to concern only knowledge claims (as the assertion that the word "know" means different things in different contexts of use). Still, at other times it is taken to regard the knowledge relation itself (as the assertion that knowledge itself has no single univocal nature). Call the former view Semantic EC, the latter view Substantive EC, and the idea that the plausibility of Semantic EC presupposes that of Substantive EC, the "Presupposition Thesis." Numerous authors argue against the Presupposition Thesis on the grounds that an understanding of the nature of knowledge is no more required to understand the meaning of knowledge assertions than an understanding of the self, for instance, is needed to understand the meaning of sentences containing "I." These authors then offer additional arguments for the same conclusion, using further comparisons between "know" and other indexicals, as well as between "know" and quantifiers, gradable and modal adjectives. Herein, I defend the Presupposition Thesis by arguing against these authors' claims (based as they are on these types of comparisons) that Semantic EC is plausible without the supposition of Substantive EC.

KEYWORDS: epistemic contextualism, knowledge, Paul Grice, Keith DeRose, Stewart Cohen

Ignoring minor endogenous disagreements, we can take epistemic contextualism (EC) to be the thesis that the standards that must be met by a knowledge claimant vary with contexts of utterance. Thus, even though knowledge claims must satisfy relatively low epistemic standards in some contexts, they must satisfy higher standards in other contexts where more remote sources of possible disinformation and error (ultimately generating skeptical scenarios) legitimately arise for consideration. Using precedent diction, we can say that contexts here are formal structures that provide values for what counts as proof, thus determining the truth values of epistemic claims. They are distinct from situations, i.e., concrete arrangements of items within which sentential utterances occur. Consequently, situations include utterances and determine contexts that generate various meanings and consequent sentential truth values that vary across contexts. A single sentence

can have different truth values at different times as a function of different contexts, which is to say, different situations in which it is uttered.

Thus construed, contextualism is generally treated as a semantic thesis that may or may not have epistemological consequences. It can concern only knowledge *claims,* or it can concern the knowledge relation itself. Let's call the view that what "knowledge" *means* depends on contextual factors "Semantic EC." Let's call the claim that what knowledge *is* depends on contextual factors "Substantive EC."[1] Let's call the claim that Semantic EC *presupposes and thus implies* Substantive EC the "Presupposition Thesis" and the denial of this position "(Epistemic Contextualist) Separatism." More specifically, Semantic EC is the view that "knowledge" discourse has an indexical status that causes the meanings and thus the truth conditions of sentences containing "know" to vary with contextually determined standards of appropriate rigor (concerning stakes, interests, etc.) Substantive EC is the view that the knowledge relation itself varies with differences in contextually determined standards of appropriate rigor (concerning stakes, interests, etc.) Finally, the Presupposition Thesis asserts that Semantic EC is only plausible on the assumption of Substantive EC. It is the view that, if true, Semantic EC provides *grounds* for Substantive EC because the contextual character of "knows" implies the contextual character of the knowledge relation as a result of presupposing it. The Presupposition Thesis thus claims that Substantive EC is a necessary condition for Semantic EC.

For illustration, consider the bearing of this account on familiar worldly skepticism. Semantic EC maintains that Moore's assertions and those of the skeptic don't conflict. Substantive EC holds that the skeptic can gain no critical traction against ordinary knowledge claims because there is no knowledge relation with a singular determinate nature at issue. The Presupposition Thesis implies that there is no acontextual, univocal meaning of "knowledge" that the skeptic can critically invoke *because* there is no singular, determinate nature that knowledge has.

In what follows, I argue for the Presupposition Thesis. Note that in doing this, I do not lobby for Semantic EC (if anything, I describe reasons to reject it). I merely maintain that the plausibility of Semantic EC presupposes that of Substantive EC. (The task of arguing against Substantive EC itself must wait for another occasion.) To this end, I propose a number of metaphysical assumptions that I take to recommend the Presupposition Thesis.

There are several things worth noting before we begin. First, I take the skeptical problematic seriously in this paper. That is, I reject any putative account of

[1] Patrick Rysiew, "Epistemic Contextualism," in *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), ed. Edward N. Zalta, URL = <https://plato.stanford.edu/archives/win2016/entries/contextualism-epistemology/>.

knowledge or knowledge discourse that doesn't at least try to say something (either positive or negative) in response to worldly skepticism (rather than merely speak past it). This has several consequences, which I describe below in due course. Also, it bears notice that I intentionally describe the type of Semantic EC with which I deal generically. Even though I invoke Keith DeRose (and a handful of other expositors) most often for exposition purposes, I aim to assume very little about Semantic EC other than what I say in the opening paragraph of this paper. To paraphrase Crispin Wright on the topic of realism, this is best because contextualism comes in so many flavors and fragrances these days that to utter the word "contextualism" is to do little more "than clear one's throat."[2] Numerous distinctions are drawn in the literature. Fortunately, for present purposes, most of these distinctions are irrelevant. It *does* matter here that contextualism regards whether certain mental states of the subject count as evidence at all.[3] For reasons I describe below, It *also* matters here that contextualism is primarily "subject-regarding" rather than "attributer-regarding" (i.e., that it concerns first-person claims about what one knows rather than third-person claims about what others know), and that the circumstances with which the truthfulness of knowledge claims vary are internalist rather than externalist. However, as far as I can see, it doesn't matter for my purposes if we describe the *mechanism* through which the shifting of epistemic standards occurs in terms of "relevant alternatives,"[4] subjunctive conditionals[5] or varying constraints on epistemic closure principles. Nor does it matter how we rank or weigh the relative significance of interests, expectations, stakes, and the like in our list of meaning-determinative factors. Finally, it doesn't matter if we take the pertinent feature of "known" beliefs to be "sensitivity," according to which one only knows p if one's belief p matches the facts in the closest ¬p-worlds[6] or "safety," according to which one knows that p only if one's belief p matches the facts in all nearby worlds,[7]

[2] Crispin Wright, *Truth and Objectivity* (Cambridge, MA: Harvard University Press, 1992), 1.

[3] Ram Neta, "S knows that P," *Noûs* 36, 4 (2002): 663–681.

[4] e.g., Alvin I. Goldman, "Discrimination and Perceptual Knowledge," *Journal of Philosophy* 73 (1976): 771-791; Fred Dretske, "Epistemic Operators," *The Journal of Philosophy* 67 (1970): 1007-1023.

[5] e.g., Robert Nozick, *Philosophical Explanations* (Cambridge, MA: Harvard University Press, 1981); Keith DeRose, "Solving the Skeptical Problem," in *Skepticism: A Contemporary Reader*, eds. Keith DeRose and Ted A. Warfield (Oxford: Oxford University Press, 1999), 183-219.

[6] e.g., Nozick, *Philosophical Explanations*.

[7] e.g., Ernest Sosa, "How to Defeat Opposition to Moore," in *Epistemology: An Anthology*, eds. Jaegwon Kim Jeremy Fantl, Matthew McGrath (Oxford: Blackwell, 2008); Timothy Williamson, *Knowledge and its Limits* (Oxford: Oxford University Press, 2000), 147.

or some other sufficiently similar property. I have little interest in the specific question of "which epistemic gear the wheel of context turns."[8]

1. The Issue: Indexical Analysis

My suspicion that Substantive EC about knowledge undergirds semantic EC is grounded in a more general conviction which, on the face of it, seems quite common-sensical: all things being equal, prior suppositions concerning the nature, specificity, and determinateness of a term's reference or denotation should inform any semantic analysis of that term. The fact that a term's *use* is prima facie amenable to a certain analysis doesn't allow us to conclude that said analysis is correct. It tells us only that this analysis is a potential candidate. Though a term's use may be subject to numerous different analyses, the analysis we settle on should be guided, though not completely determined, by our beliefs about the natures of these terms' denotations and what these natures tell us about how these terms effect reference. To do this is to guide our views about language, in part, by our views about ontology. To those who think that such a procedure has things backward and that we should base ontological commitment solely on linguistic practice, I commend an account of things on which language is merely one naturalistic phenomenon among others, and thus part of a world of language-independent facts which it may only imperfectly describe. On the supposition that the primary role of assertive language is to represent reality, reality constrains language, not the other way around. To contend with Quinean pithiness, "ontology recapitulates philology" only to the extent that philology gets things right in the first place.

The claim that ontology recapitulates philology is weaker than the Separatist doctrine that there is ***no*** essential connection of any sort between semantic analysis and metaphysical presupposition. Quine's view, taken at face value, is a claim about the order of semantic and ontological considerations. Separatism, on the other hand, is the view that Semantic and Substantive EC enterprises are as apples are to oranges. Separatism is a view succinctly expressed by Patrick Rysiew.[9] As he tells it, Semantic

[8] Jonathan Shaeffer, "What Shifts? Thresholds, Standards, or Alternatives?," in *Contextualism in Philosophy: Knowledge, Meaning, and Truth*, eds. Gerhard Preyer and Georg Peter (New York: Oxford University Press, 2005), 115. Personally, I believe with Kornblith that so-called "safety" accounts should be modified to include a "sensitivity" element. Sosa himself offers his safety account as only a first approximation, and writes, "What is required for a belief to be safe is not just that it would be held only if true, but rather that it be based on a reliable indication." [Sosa, "How to Defeat," 286] Such an idea of being "based on a reliable indication" suggests a notion of "responsiveness," which itself smacks of sensitivity.

[9] See Rysiew, "Epistemic Contextualism."

EC is a thesis about the truth conditions of knowledge sentences in specific contexts (i.e., the propositions expressed by utterances thereof). In contrast, Substantive EC is a thesis about the knowledge relation itself. As one more contextual term amongst others, Semantic EC's accuracy as a correct semantic analysis of "know" can, therefore, receive support from comparisons with other ordinary language indexicals, such as "here," "now," and "I." The truth conditions of a tokening of the sentence "I hunger for pie" clearly depend on contextual features concerning the speaker's identity, Rysiew concedes, since these features precisely determine which proposition is expressed (who's hungry, you or me?) However, the mere fact that "I" is token reflexive, Rysiew insists, tells us nothing about the nature of the speaker or about the nature of the self more generally. Similarly, he suggests, the contextuality of "x knows that p" tells us nothing about the nature of knowledge. The semantic analyst is concerned to tell us only when the statement "x knows that p" is truthfully asserted. In contrast, the epistemologist is concerned to tell us when x knows that p.[10]

Thus, we are told, the question of what a subject knows is different from the question of what knowledge claims are true of that subject, allowing contextualists to refrain from issuing first-order judgments about knowledge. As Geoff Pynn writes, to truthfully say of a subject that she knows is to claim, in effect, that she satisfies the epistemic standards in place in [the applicable] context. The case for contextualism, Pynn tells us, doesn't involve "some intuitive judgment that [a subject] meets or doesn't meet the epistemic standards in place in the context of a philosophical discussion about knowledge or knowledge claims. Instead, it depends on the judgment that subjects' knowledge claims, as made in their imagined contexts, are true."[11]

Most contextualists express sympathy with this idea. Both DeRose and Cohen, for instance, claim to be concerned with whether speakers use "know" correctly, with whether they "speak truly," not with whether said speakers actually know. DeRose claims that he "find[s] skepticism persuasive and [merely] wants to explain the persuasiveness of the skeptic's attack."[12] Cohen maintains that contextualism "preserves our belief that we know things" while "explaining the undeniable appeal of skeptical argument."[13]

---

[10] See Rysiew, "Epistemic contextualism."

[11] Geoff Pynn, "The Intuitive Basis for Contextualism," in *Routledge Handbook of Epistemic Context*ualism (Routledge Handbooks in Philosophy) (Oxford: Taylor and Francis, Kindle Edition, 2017), 34.

[12] Keith DeRose, "Assertion, Knowledge, and Context," *Philosophical Review* 111 (2002), 168.

[13] Stewart Cohen, "Contextualism and Skepticism," *Philosophical Issues* 10 (2000): 100.

2. Anomalousness of an Indexical Analysis of "Know"

My argument herein is that DeRose's and Cohen's claims above are mistaken. To provide an adequate account of "knowledge" discourse to the end of "explain[ing] the persuasiveness or undeniable appeal of the skeptic's attack," we automatically disallow ourselves from retaining the distinction that most contextualists posit between Semantic and Substantive issues.[14] However, fleshing out this argument requires the length of this paper. Let's start with first impressions. What should we make of the sort of Separatist argument against the Presupposition Thesis described above?

A recurrent contextualist defense of such Separatism invokes the distinction between subject and attributer stances and the idea that proponents of the Presupposition Thesis suffer from a "levels confusion" between the first-person subject's and third-person contributor's knowledge. Conversational propriety (a factor we consider at greater length below) partly determines whether an attributer can truthfully describe a subject as "knowing." The reason is that such factors affect the content of a third-person attributer's claim without affecting the subject's own epistemic state.

It is critical to note that this move is inapplicable in the present dialectical context. As noted, I take contextualism to be subject-regarding (rather than attributer-regarding), and the circumstances with which the truthfulness of knowledge claims vary to be internalist rather than externalist. These assumptions are not *ad hoc* on my part. As will be remembered, I take skeptical challenges to empirical knowledge seriously. Thus, I take the adequacy of contextualism to depend, in part, on its ability to say *something* (either positive or negative) in response to the skeptic's concerns. Attributer-regarding forms of Semantic EC are unable to do this. Deep skeptical uneasiness cannot even be articulated from the stance of attributers, given that their judgments about the truth values of subjects' knowledge claims necessarily arise from within a naturalistic background of presupposition that they share with these subjects themselves. One cannot purport to offer an attributer account in third-person terms which explains "the persuasiveness of the skeptic's attack" or "the undeniable appeal of skeptical argument" because such accounts manifest a deep misunderstanding of the stance from which the skeptic poses his challenge, and thus a deep misunderstanding of what the skeptic's attack and argument are. Neither can the contextualist, *qua* contextualist, purport to say anything of interest about skepticism when his initial externalist presuppositions alone do most of the anti-skeptical work before

[14] Keith DeRose, "Assertion, Knowledge, and Context," *Philosophical Review* 111 (2002): 168.

contextualist insights even arrive on the scene.[15] For these reasons, I write about *first-person* "knowledge" and knowledge in the remainder of this paper.

This said, it is not at all clear what the above-described distinction between the Semantic and Substantive enterprises is supposed to be, at least when mind-independent truth (instead of "warranted assertability," "conversational appropriateness," or some such alternative feature [which Pynn invokes under the label "propriety"], directed at something other than truth) is at stake.[16] If we assume, as have I, that ontological supposition should inform semantic analysis, then the kind of distinction between object linguistic and meta linguistic investigation which Rysiew and Pynn propose applies in connection with syntax much more clearly than it applies in connection with semantics. For the disquotationalist, "x knows that p" is true iff x knows that p, *even if* we take the content of "know" to be determined by the situation in which its utterance occurs. Let's subscript both "knows" and the knowledge relation of this context ("know$_1$", knowledge$_1$) to make it clear that for Semantic EC, both the sense of the term and identity of that term's referent are context-dependent. Even when we do this, our T-Sentence dictates that we only understand what it is for the truth predicate to apply in the metalanguage by referencing the objects and relations picked out in the object language. Understanding when "I know$_1$ that p" is truthfully asserted requires that I understand what it is for me to know$_1$ that p. However, to understand this, I must surely ask myself what kind of relation "know$_1$" (as opposed, say, to "know$_2$") denotes. To claim Semantic EC is to claim that knowledge is, at most, a disjunctive relation. Thus, Semantic EC makes a substantive claim about the knowledge relation itself, not just knowledge assertions. Specifically, it maintains that there is no knowledge relation, *simpliciter*, but instead numerous different knowledge relations somehow appropriate to different contexts of inquiry. But such a position, I submit, is just Substantive EC.

This is hasty and thus suspect. So, let's take things more slowly. We can begin by briefly canvassing contemporary trends that have led to EC. These trends are numerous: Malcolm's distinction between strong and weak knowledge[17] and the epistemic pluralism suggested by Wittgenstein's talk of different language games designed for different purposes;[18] relevant alternatives approaches to understanding

[15] Kornblith makes much the same point concerning the role that a prior externalism plays in contextualist responses to skepticism (Hilary Kornblith, "The Contextualist Evasion of Epistemology," *Philosophical Issues* 10 (2000): 24–32).

[16] See Kornblith, "Contextualist Evasion."

[17] Norman Malcolm, "Knowledge and Belief," *Mind* 61, 242 (1953): 178-189.

[18] Ludwig Wittgenstein, *Philosophical Investigations* (New York: Macmillan, 1953).

knowledge claims, as anticipated by Austin[19] and more explicitly articulated by Dretske[20] and Goldman;[21] the social epistemology movement championed by the likes of Annis[22] and Rorty;[23] and attempts to deal with Gettier cases with the idea that variability in expectations calls for variation in epistemic standards (Pollock, who credits Harman, who in turn credits Sosa).[24]

It is not always clear that these precedent quasi-contextualist treatments are offered as exclusively semantic theses, that is, as theses concerning "knowledge" rather than knowledge. But explicitly Semantic EC seems to come into its own with the advent of the sort of indexical analysis articulated by Rysiew and Pynn above. So, to give Separatism, the idea that Semantic EC is independent of Substantive EC, its due, we would do well to examine some of the contextual terms that are most commonly invoked as precedents to support EC as a distinctively Semantic thesis. If the natures of the items that these terms regard and the ways that these terms affect reference seem significantly different than in the case of "know," this should prove illustrative. How does "know" compare to the sorts of items that contextualists commonly invoke as precedent indexical contextualist terms on whose treatment Semantic EC is allegedly modeled.

Indexical analyses were first offered by Kaplan of "pure indexicals" (terms like "I," which automatically pick out referents irrespective of a speaker's intentions) and "true demonstratives" (terms like "he" which do require an accompanying speaker's intentions or gestures).[25,26] I suggest that the first thing to note is that the indexical analysis of "know" is *prima facie* contentious in a way that corresponding contextualist analyses of these other terms are not. Epistemology displays a long tradition in which "knowledge" is taken to refer to an invariant, unchanging, specific, and determinate relation between claimants and the world, mediated by

---

[19] J. L. Austin, "Other Minds," *Proceedings of the Aristotelian Society*, Supplementary Volume 20: 148–187. 113; cf. *ibid.*, 88.

[20] See Dretske, "Epistemic Operators."

[21] See Goldman, "Discrimination."

[22] David Annis, "A Contextualist Theory of Epistemic Justification," *American Philosophical Quarterly* 15 (1978), 213-219.

[23] Richard Rorty, *Philosophy and the Mirror of Nature* (Princeton: Princeton University Press. 1979).

[24] John Pollock, *Contemporary Theories of Knowledge* (Savage, MD: Rowman & Littlefield Publishers, 1986); Gilbert Harman, "Knowledge, Inference, and Explanation," *American Philosophical Quarterly* 5-3 (1968): 164 -173.

[25] David Kaplan, "Demonstratives," in *Themes from Kaplan*, eds. Joseph Almog, John Perry, and Howard Wettstein (New York: Oxford University Press, 1989): 481–563.

[26] David Kaplan, "Afterthoughts," in *Themes from Kaplan*, 565-416.

evidence. In itself, this fact doesn't speak for or against contextualism, as the traditional invariantist view might simply be misguided. However, it does speak to a distinct difference between "knowledge" talk and the various pronouns, adverbs, and adjectives to which indexical analysis is otherwise typically applied. Semantic EC's advocates present their view as a recent discovery about "know." The likes of "I," "this," "now," "here," "tomorrow," however, were not discovered to be contextual: they were introduced into English to serve indexical functions from the very outset. This fact explains why philosophers have never dined out on the issue of these terms' indexical status, and it grounds the wisdom of Schiffer's observation that it is a general linguistic truth that speakers automatically recognize implicitly relative predications when they occur.[27]

Moreover, when construed as the name of a relation, "know" is very unlike these other examples. "I know that p," where "know" is construed as a two-place relation term, is structurally akin to "I hunger for pie." However, "hunger for," not "I," stands in a predicate position parallel to the position of "know that." And while it may be true (paraphrasing Rysiew) that we need not understand anything about the nature of the self when specifying the conditions under which we truly assert "I hunger for pie," we surely *do* need to understand something about the nature of hunger for this assertion to have content for us. Similarly, while it is true that we may need not understand anything about the nature of the specious present or simultaneity when specifying the conditions under which "Fred is taking a test now" is true, we surely do need to understand something about the nature of test-taking. Similarly, while it is true that we need not understand anything about the nature of temporal passage when we assert, "Fred is driving to Cincinnati tomorrow," we surely do need to understand something about the nature of driving. And while it may be true that we need not understand anything about the character of the excursionist in question when we assert "He is traveling tomorrow" (perhaps with accompanying ostensive gestures), we surely do need to understand what it is to travel.

The reason for these differences is that to exhibit hunger, take a test, drive a car, or travel is to display a determinate nature or quality satisfying what we might call "non-conversational" criteria, which is to say that to have these properties and relations is a function of something other than satisfying various indicative conversational conventions. In the indexical use of "I," allusion to myself is affected by a direct referential performance devoid of descriptive content. However, reference to "hunger" is not directly affected in this way because the exhibition of

[27] Stephen Schiffer, "Contextualist Solutions to Scepticism," *Proceedings of the Aristotelian Society* 96 1996): 317-333.

hunger is identified through something more than the exercise of various verbal moves supplemented, perhaps, by appropriate motions. "Hunger" designates hunger by satisfying criteria that do or don't obtain irrespective of social performance. For convenience, let's say that a property or relation of this sort has "non-conversational specific criterial determinacy (or "NCS Criterial Determinacy" for short)." Further, we can say that terms like "instantiates hunger," "takes a test," and "drives a car" either do or don't apply when they stand for "non-conversational specific criterial determinants" (or "NCS Criterial Determinants").

I am not claiming here that terms with NCS Criterial Determinacy designate specific items in a way that simple indexicals and pure demonstratives do not. On particular occasions of use, "I" obviously refers to a particular person. Neither am I claiming that indexical and demonstrative terms lack *functional* dictionary definitions concerned to describe what they *do*: "I" clearly means something along the lines of "the first-person singular used by a speaker to refer to himself or herself." Finally, I am *not* claiming here that the "is hungry" relation, unlike "I" (across *all* occasions of use), has a singular nature (though I presume it does).

The only claim I am pressing here is the following: "it" (or "now" or "she," or whatever), unlike "being hungry," affects reference without the individuating mediation of descriptive content, the concern of which is to describe a nature. This is why we need to understand the meaning of "hunger" to understand sentences in which it occurs in a way that we need not understand "I" to understand sentences in which it occurs. It is also why there is a question of getting things right or wrong in the former cases that doesn't arise in the latter sorts of cases. In these latter cases, the terms apply (even when the sentences in which they occur are false) immediately by virtue of agreed-upon performative conventions. When said performances occur, these terms enjoy immediate application without regard for whether various other non-performative criterial conditions have been satisfied. Once again, this is symptomized by the fact alluded to above. "I," "this," "now," "here," "tomorrow" were not discovered to be contextual: they were introduced into English to serve indexical functions from the very outset. No one ever imagined that the generic "this" was anything other than a term stipulated to perform an indexical function, designed to have variable reference depending upon what one intends or where one happens to point. Nothing had to be known about the range and natures of this term's denotations for it to enjoy immediate use.

"Knowledge," on the other hand, is at least a *candidate* denoter of a relation with such a nature. I submit that this should make us more reluctant to assume that it doesn't denote a relation with such a nature, given that questions about its nature (unlike in these other cases) have a long-standing philosophical pedigree. However,

my present concern is not to press this claim. My immediate concern is to advance the claim that *if* we assimilate "know" with unproblematically indexical terms, *then* we probably commit ourselves to a contentious view about knowledge itself (i.e., that it lacks an NCS Criterial Determinant nature). And this view, to repeat, is just Substantive EC.

Quantifiers, which have also been proffered as candidate precedents for a contextualist analysis of "know," are similar to indexicals and demonstratives in this respect.[28] Even though "all," for instance, has a functional dictionary definition (along the lines of "used to refer to the whole quantity or extent of a particular group or thing") which tells us what the word does, this is not a definition that gives the word a specific, determinate denotation in the absence of various verbal moves supplemented, perhaps, by appropriate nods and gestures. Or, more precisely, it doesn't do this unless augmented by additional phrases (e.g., "(All) people occupying classroom six in the Humanities building of the University of Woolamaloo campus") which do denote NCS Criteria Determinants.

3. Gradable and Modal Adjectives

Pure indexicals, true demonstratives, and quantifiers are not the only terms that have been served up as precedents for a contextualist account of "know." Other terms, subject perhaps to less blatantly indexical analysis, include gradable adjectives, for instance, "warm," "tall," "large," "heavy," and "fast."[29,30] These terms are clearly context-sensitive: the truth value of "My uncle is tall" varies across contexts as a function of which comparison classes prove salient. However, even though these terms have more substantive content than the pure indexicals and true demonstratives considered above, they differ from "knows" in much the same way: they aren't candidate denoters of objects or relations with specific determinate natures captured by non-conversational, non-performative criteria. The reason is that, as property terms go, they shouldn't be taken to denote at all. We don't really believe that there is a property of being warm (full stop) any more than we believe that there is a property of being rich (full stop). But this is because "warm," "rich," "tall," and the like are inherently comparative: a subject can at most be richer (or warmer or taller) than another. Think of the problems Plato encountered when he

[28] Jonathan Ichikawa, "Quantifiers and Epistemic Contextualism," *Philosophical Studies* 155, 3 (2011): 383–398.

[29] Stewart Cohen, "Contextualism, Skepticism, and the Structure of Reasons," in *Philosophical Perspectives* 13, Epistemology (1999): 57-89.

[30] Keith DeRose, *The Case for Contextualism: Knowledge, Skepticism, and Context*, Vol. 1 (Oxford: Oxford University Press, 2009), 185.

considered taking such terms to denote in the concrete world. The earth became a contradictory mess as the selfsame objects became warm and cold, tall and short, large and small. It was primarily for this reason that he eschewed the empirical realm as a hall of mirrors possessed of incomplete reality (not that an appeal to forms did any good). The denotation of "warm" does not have NCS Criterial Determinacy. The denotation of "warmth" does, but *only* if understood as a quantifiable physical magnitude positioned along a linear scale of gradation. However, understanding warmth in this way is to understand it as physical heat or temperature, thus removing its contextual features entirely.

To be clear, while "warm" doesn't denote a property with a specific determinate nature, "is warmer than" presumably does. Specifically, it denotes a two-place comparative relation possessive of NCS Criterial Determinacy. However, this phrase and others like it do not provide us with a precedent for thinking that Semantic EC is true, as there is no compelling sense in which "is warmer than" is contextual. While the truth-values of sentences arrived at by substitution within [x is warmer than y] vary with values for x and y, this is not due to the contextuality of "warmer than," since said contextuality regards issues of reference and denotation, not truth. Undoubtedly, the truth values of sentences of the form "x is warmer than y" do vary with such substitution instances. However, the meaning of "warmer than" remains constant across all its applications: it refers to a single relation, the nature of which is clarified by examples of the true sentences in which it occurs. In other words, "warmer than" purports to designate something with NCS Criterial Determinacy. Or it is at least a candidate example of such in a way that "warm" is not.

In addition to gradable adjectives, other terms have been offered as candidate precedents for a contextual analysis of "knows," modal words like "necessarily" and the subjunctive conditional connective "if… then."[31] The idea here is simple enough: depending on one's interests and purposes, one might mean, say, logical rather than nomological necessity or implication. However, these are telling examples not because they fail as candidate denoters of NCS Criterial Determinacy, but because their indexical status is dubious to begin with. Unlike the gradable terms described above (e.g., "large," "rich"), these are not ordinary language adjectives subject to loose and unthinking use. Instead, they demarcate technical distinctions between types of possibility and necessity, distinctions originally introduced into English to serve very specific functions. Philosophers and mathematicians don't unknowingly mean "logical necessity" as a function of conversational contingencies. These phrases

[31] Jonathan Ichikawa, "Introduction: What is Epistemic Contextualism?," in *The Routledge Handbook of Epistemic Contextualism* (Oxford: Taylor and Francis, Kindle Edition, 2011).

mean as a function of speakers' determinate concrete intentions. Given this, I submit, we have no more reason to regard "possible" or "implies" as contextual than we have to view them as merely ambiguous. We have no more reason to treat them as contextual than we have to treat "still" ("immobile" or "quiet") as contextual.

4. Conversational Propriety Conditions

Note again both what my overall claim so far is and is not. I am not contending that the terms to which Semantic EC often compare "know" either are or are not indexical. I am claiming that *if we take* these other cases as illustrative precedents for a contextualist analysis of "knows," then we express sympathy with the view that "knowledge," like these other terms, does not refer to an NCS Criterial Determinant. But this view, I have maintained, is just Substantive EC; it is the claim that knowledge is not a single subject matter with a unitary character. The primary difference between "know" and these other terms is that this anti-realist consequence is only contentious in the case of "know." It is contentious because, unlike these other cases, where we never imagined that our use requires us to affect reference by understanding substantive concepts capturing real properties or relations, this is at least an open question in the case of "know."

One last time, I have argued for this openness in several ways. For one thing, the grammatical character of "know" differs significantly from that of indexicals and demonstratives. We cannot go wrong in applying, e.g., the concept "I," in the way that we might go wrong in using the concept "know" because there is no corresponding way in which we could go wrong. This is because "I" effects reference directly rather than through intermediate descriptive content of the sort that might pick out a kind. For another thing, the contextual character of "know" is routinely proffered by contextualists as a discovery rather than a stipulation, as reflected in the fact that questions about the nature of knowledge have a long-standing philosophical history. Thus, to assert that knowledge has no single, unitary character is to make a contentious and, therefore, substantive claim. In this, it is unlike the claim that the referent of "I," in its various contexts of occurrence, has no single, unitary nature. This latter claim is a nothing burger. This former claim is an assertion of singular epistemological significance.

Perhaps we can circumvent this joining at the hip of Semantic and Substantive EC by de-emphasizing the aim of providing truth conditions for knowledge assertions and emphasizing the aim of offering "propriety" conditions instead as their determinants of meaning. Pynn distinguishes between these two projects when describing the challenges faced by invariantist accounts, which, he claims, fail to accord with our ordinary intuitions concerning when our knowledge assertions are

veridical or appropriate.[32] Pynn explicitly distinguishes between the two criteria for understanding knowledge attribution sentences described above as the "truth challenge," which requires that the invariantist explain how a knowledge sentence that seems true (e.g., a prosaic assertion about a subject's knowledge of her own handedness offered in a Moorean spirit) really isn't, and the "propriety challenge," which requires that the invariantist explain how a seemingly improper truth attribution about a knowledge claim (e.g., a negative skeptical pronouncement about a subject's knowledge of his own handedness, offered in a Cartesian spirit) is actually "proper." Pynn takes care to distinguish these as two *distinct* challenges and to say that conversational propriety is a measure of something other than truth.

Here "appropriateness" commonly regards the pragmatic consequences of utterance. The idea is that "truth conditions" are assisted by "conversational propriety" conditions to determine meaning. Suppose we give due credit to this aspect of contextualist accounts. Does it, by de-emphasizing truth conditions, somehow yield a Semantic EC which avoids commitment to Substantive EC?

I will not pause here to consider DeRose's critique of specific so-called "warranted assertability arguments." This critique is pertinent to the issues at hand. However, it calls for a lengthy discussion that I can only offer here in glancing by way of extended endnote.[i] Instead, let me argue from a different direction for a claim that even when we turn from truth to warranted assertability conditions in the manner suggested above, Semantic and Substantive EC still remain thick as thieves, as the former recommends the latter. To the extent that we regard conversational appropriateness, understood as something other than truth, as determinative of the meaning of knowledge attributions, we automatically regard the *subject matter* of knowledge talk to possess second-rate ontological standing, as something less than a real, non-conventional relation. In effect, we deny that knowledge has NCS Determinacy. That is, we end up embracing Substantive EC anyway.

In one respect, this might seem like an odd claim. The conversational propriety of a term's use, after all, looks to be an entirely semantic or pragmatic affair, a measure of nothing more than the degree to which various conversational norms have been satisfied. Thus, we might ask, shouldn't measures of conversational propriety be a purely linguistic matter, implying nothing about the alleged referents of the terms at issue? This diagnosis, however, doesn't do full justice to the topic at hand. For, we typically think that semantic analyses in terms of conversational propriety conditions, understood as something other than truth conditions, are appropriate in cases of concepts that fail to denote fully "real" properties existing as a function of what Searle calls "brute" or "non-institutional facts." Searle's definition

[32] See Pynn, "Intuitive Basis."

of "institutional facts" proceeds in several stages. Take ontologically subjective features to be features of mind-dependent objects in the sense that their existence depends on being experienced by subjects (e.g., phenomenal redness as opposed to electric charge). Take "relative to the observer intentionality" features of objects to be those which fail to exist independently of observers' representations (e.g., being a screwdriver as opposed to being a metal rod with a narrowed blade at one end and a widened casing at the other). Take "agentive functions" to be features of objects which we intentionally assign rather than simply observe or discover (e.g., a rock's serving as a paperweight as opposed to a heart's pumping blood). Take "constitutive rules" to be maxims that define activities into existence (e.g., the rules of chess as opposed to rules of the road, which simply "regulate" the existentially autonomous activity of driving). We can now say that constitutive rules allow people to create social facts by human agreement. It may be an ontologically objective fact that I consume proteins, fats, and carbohydrates without utensils on some particular occasion, but it is merely an institutional or social fact that I display bad manners by eating lasagna with my hands at a black-tie dinner.[33]

To harp further on rules of decorum, consider specifically the etiquette of verbal exchange. We do not think of an act's being socially well-behaved as a genuine, non-institutional property. It is for this reason that we might be inclined to think of "S politely demurred from replying in kind" as a sentence to be understood solely in terms of its local norms of acceptance because it is a case in which truth conditions consist in nothing above and beyond conditions of conversational appropriateness. And though it need not generally be the case that "non-institutional" reality-status and NCS Determinacy march in lockstep (the categories are not coextensive), it seems reasonable to equate them here to the extent that the institutional reality of conversational politeness indexes the meaning of "politeness" to the local ways and customs of the settings in which sentences attributing such propriety to speakers are uttered.

Again, it is important not to misunderstand my point here. My claim is not that "knows" (as a verb) designates a "real" property or relation in a way that "politely" (as an adverb) does not. For all I've said, it could easily be the case that neither "knowledge" nor "politeness" denotes "real" non-institutional properties with NCS Criterial Determinacy. My present point is different: Semantic EC analysis that invokes the evidential authority of "conversational propriety," conceived as something distinct from truth, implicitly suggests a view on which "knowledge" possesses mere institutional or social reality devoid of an intrinsic nature, the lack of which prevents it from enjoying NCS Criterial Determinacy. Thus, the

[33] John R. Searle, *The Construction of Social Reality* (New York: Free Press, 1997).

contextualist's reliance on such factors of "conversational propriety" presupposes a prior commitment to a view of knowledge as a disjunctive relation devoid of a single univocal nature. And such a view of knowledge, we have noted, is just Substantive EC.

I am speaking above about the case in which "conversational propriety" conditions are viewed as something distinct from truth conditions. We asked if such an approach to semantic analysis might help us avoid the problem we encounter when we consider truth conditions alone. I have maintained that it doesn't. On reflection, however, I think it clear not merely that this approach leads nowhere, but that this approach isn't even an option. For I fail to see how the evidence Pynn invokes to support his "intuitions" about when knowledge attributions are true differs in any significant way from the evidence he invokes to support his intuitions that such knowledge attributions satisfy rules of conversational propriety. The function of declarative sentences is to assert putative facts (or at least present them for contemplation or consideration, as in fiction). Given this, the following is hardly a surprise: to maintain that it is odd to say that a declarative knowledge attribution sentence is true is to asset little more than that said sentence is odd to say.

The most canonical source to consult when asking about the conversational propriety of an utterance, of course, is Grice's list of conversational maxims, his rules of quantity, quality, relation, and relevance.[34] The rule of quantity maintains that speakers should be appropriately informative, providing all and only as much information as needed. The rule of quality holds that speakers should avoid giving false or unsupported information. The rule of relation states that speakers should only say things that are pertinent to the discussion at hand. Finally, the rule of manner commends clarity, brevity, and the avoidance of obscurity of ambiguity. Of these, only the rule of quality may seem, on first inspection, to explicitly regard veridicality. Thus, it may seem that only the rule of quality is a rule according to which an utterance's apparent conversational oddity is tantamount to said utterance's seeming false. On closer examination, however, I submit that the other three maxims express mostly the same concern: maximizing true (or, at least, minimizing false) belief.

Consider the rule of quantity, which dictates that utterances be appropriately informative, providing all and only as much information as is required in the pertinent circumstances. The aim of this maxim is mainly to head off false utterances at the pass. Imagine that I ask a passenger carpooler if there is a gas station nearby, and he replies, "Yes, there is," knowing full well that said station is closed due to the

[34] Paul H. Grice, "Logic and conversation," in *Studies in the Way of Words* (Harvard; Harvard University Press, 1991).

recent death of its manager in a freak and tragic blimp accident. His utterance is not odd in the way that it would be odd if he were to suddenly curse my mother. It is odd because he has provided false information to the question which he knows full well I am asking (i.e., "Can I buy gasoline around here?")

Consider the rule of relation, which maintains that speakers should only say things that are pertinent to the discussion at hand. Imagine the situation above, but with the following difference: my interlocutor responds by telling me that gasoline is plentiful in our county (again, while knowing full well that the only local station is closed due to the aforementioned tragedy), perhaps adding that in our state commercial gasoline is required by law to be diluted by 10% ethanol. Ignoring violations of etiquette or expectation (suppose this time he punctuates his sentence with a complicated choreography of belching and tap dancing), the oddness of his response is that it is misleading. He is very simply not answering the question which he knows full well I am asking, *and* he is inciting inaccurate inferences on my part by prompting me to misread the import of his words.

Consider the rule of manner, which recommends clarity, brevity, order, and a lack of obscurity and ambiguity. Imagine my interlocutor replying with the claim, "I have heard stories from many a weary traveler about a retail enterprise in these parts that sells fuel to those who hanker for commercial exchange." Suppose that he is correct in what he says, but only by virtue of the fact (of which he is fully aware) that a local convenience store dispenses butane tanks for barbeques. The concern for clarity, brevity, order, and non-obscurity are partly stylistic here. However, the primary concern at play is still that of accuracy. Clarity, brevity, order, and comprehensiveness expedite the utterance of determinate, unambiguous, and thus contentful propositions (the only sorts of things possessed of truth values). And the injunction against ambiguity prevents the formation of false inferences on the part of hearers.

This should hardly come as a surprise. Granted, judgments about appropriateness, for Grice, are also driven by the like of our sensitivities to the demands of rational co-operation with our conversational partners. Granted, he famously presents his conversational maxims as corollaries of a cooperative principle designed to ensure that participants in verbal exchanges make "conversational contributions such as are required, at the stages at which they occur, by the accepted purpose or direction of the talk exchange." Granted, he concedes that there are numerous "other maxims" (aesthetic, social or moral in character), such as 'Be polite!,' which are also generally observed by participants and can also generate non-conventional implicatures." However, the conversational maxims of concern to him are specially connected with "the particular purposes that talk exchange is adapted

to serve and is primarily employed to serve."[35] Since the function of declarative sentences is to assert putative facts (or to at least present them for contemplation or consideration, as in the case of fiction), it is to be expected that the oddity of asserting that a knowledge assertion sentence is true amounts to little more than the oddity of asserting said sentence at all. Consequently, I find the very distinction between Pynn's "truth challenge" and "propriety challenge" puzzling. Given the "purpose and direction" of declarative "talk exchange," the propriety challenge is, for the most part, identical with the truth challenge.

5. Conclusions

I have made two fundamental assumptions above. First, I have assumed that contextualism should have something to say (either positive or negative) about worldly skepticism, and that, to do this, it must be construed as internalist and subject (rather than attributor) centered. Second, I have assumed that all things being equal, prior suppositions concerning the nature, specificity, and determinateness of a term's reference or denotation should inform (even if they do not wholly determine) the semantic analysis of said term. From these assumptions, I have argued (in reverse order) for a number of conclusions. First, I have argued that the very distinction between conversational propriety and truth conditions is suspect, as it becomes a distinction without a difference, given the essential point and purpose of assertion. Second, I have argued that reliance upon "conversational propriety" conditions for the semantic analysis of knowledge talk, when *accepted* as something distinct from truth conditions, suggests a commitment to the view that facts about knowledge are merely "institutional," and thus suppose a prior acceptance of the view that knowledge is devoid of a mind-independent nature (i.e., which is tantamount to accepting Substantive EC). Third, I have argued that a contextualist understanding of "know" in terms of Tarskian truth conditions presupposes a prior acceptance of Substantive EC because it fosters an understanding of knowledge as a disjunctive relation devoid of NCS Criterial Determinacy. Finally, I have argued that attempts to understand the alleged contextuality of "know" on the model of simple indexicals also presupposes Substantive EC, as such indexicals also fail to denote NCS Criterial Determinants. My overall conclusion is the following: if we think that semantic analysis should be informed by ontological commitment, we have good reason to believe that the presupposition of Substantive EC gives Semantic EC its force and motivation. The former is necessary for the latter, and Semantic EC is only as plausible as Substantive EC. "Knowledge" refers to

[35] See Grice, "Logic and Conversation," 28.

knowledge. Thus, we have no reason to think that "knowledge" is an equivocal term unless we suppose that knowledge is a disjunctive relation devoid of a unitary nature. This paper is preliminary to a larger project. Suppose we accept the preceding conclusion that Semantic EC is only as plausible as Substantive EC because the former position is given force and motivation by the latter. Remember that I do not argue above that Semantic EC is false. I argue merely that its plausibility cannot be separated from that of Substantive EC. Given this connection, it is now reasonable to ask if the evidence from use, which is commonly cited in support of Semantic EC, gives us reason to believe the doctrine of Substantive EC upon which the plausibility of Semantic EC depends. But, as I have noted, this is a project which must wait for another occasion.

[i] DeRose does offer more direct arguments against warranted assertability accounts of "knowledge." However, my responses to these arguments are too lengthy to be dealt with in the main body of this text. Thus, I am forced to address them here.

Let's first consider the epistemic case, about which so much more has been written. In part, DeRose recommends this construal on the basis of his alleged demonstration that invariantist analyses fail, for independent reasons, to explain away our actual patterns of knowledge ascription in a convincing fashion. Contextualism would win support, in this case, for lack of competing alternatives. It would win, in effect, by being the only game in town. On DeRose's telling, the most widely advocated mechanism for explaining away such patterns of usage is that of "warranted assertability." However, DeRose argues that no such account could ever explain the observed mechanics of epistemic attribution. (See DeRose, "Solving the Skeptical Problem"). DeRose's reasoning here is largely comparative: "know," he argues, proves a poor candidate for such a diagnosis when compared to other terms to which such warranted assertability maneuvers (i.e., WAMs) clearly apply. More specifically, DeRose maintains, there are several conditions governing the proper application of WAMs, constraints which knowledge attribution fails to satisfy.

On DeRose's telling, there are essentially four such constraints, reordered here for convenience. The first constraint is that WAM-amenable assertions (e.g. "It is possible that P," uttered in circumstances wherein one knows very well that P) share apparent truth-values with their contradictories (e.g. "It is not possible that P"). The second constraint is that true implicatures cannot substantively modify a person's attitudes toward false assertions. That is, one may not judge an assertion to be true simply because one judges its conversational implicature to be true. The third constraint is that WAM-amenable statements (e.g., "It is possible that P") seem false as a result of conversationally implying other sentences that are false (e.g., "I don't know that P). The fourth constraint is that such conversational implicatures occur in accordance with general and systematic (rather than particular and *ad hoc*) rules of conversational implicature. Together, these four conditions on legitimate WAMs are supposed to block the epistemic invariantist's use of this strategy, forcing us to accept the contextualist's alternative proposal. The conditions fail to do this, however. To see why, let's consider each in tum.

The first constraint on WAM-amenable assertions, again, is that they share apparent truth-values with their contradictories. Thus, if P seems false by virtue of a false implicature in a given conversational context, so must not-P. DeRose seemingly presumes here that it is only when this occurs that we have good reason to suspect the occurrence of a genuine semantic pathology, as it were, requiring an exotic WAM-type explanation rather than a much more mundane account in terms of context-variant meaning. Arguments using this first constraint fail to disqualify knowledge WAMs, however, since the invariantist can easily describe the relevant cases as ones in which the constraint is met. Consider an example of philosophical dialectic that fairly represents the manner in which skeptical scenarios are actually introduced. A student is asked, in a prosaic conversational setting, if she knows that she has hands (P). She replies in the affirmative, only to be confronted by a skeptical scenario (e.g., brains in vats or whatnot). How should we view what transpires next? On DeRose's account, the situation must be one in which the meaning of the question "Do you know that P?" changes from one context to the next. In the prosaic setting, the looser standards in play render her response "I know that P" seemingly true, and her alternative possible response "I don't know that P" seemingly false. In the skeptical context, the higher standards in play render her response "I know that P" seemingly false and the alternative response "I don't know that P" seemingly true. Thus, if we assume, with the contextualist, that the meaning of "know" shifts with conversational context, DeRose is correct: there is no context in which both "I know that P" and "I don't know that P" strike us as simultaneously true (or false). However, the invariantist can easily describe this case as one in which we are so inclined to view these claims as simultaneously true (or false). On this alternative description, the ordinary claim to know that P seems true until challenged by the citation of a skeptical scenario, at which point its appearance of truth fades and one is forced to recognize its falsity. The assertion of knowledge remains constant across contexts; all that changes is the depth of reflection with which we challenge it. The central difference between this account and DeRose's is the following: because content is no longer assumed to be determined by context, the seeming truth-values of claims can be contrasted and compared across contexts and not merely within them. The seeming falsity of a person's claim to know that P (in the face of skeptical scenarios) can be contrasted and compared with the seeming falsity of a person's claim to not know that P (in prosaic settings). Thus, DeRose's use of his first constraint against invariantist accounts of knowledge WAMs fails by virtue of its essential circularity: his argument is only effective on the prior assumption that contextualism is true.

The second constraint, again, is that true implicatures cannot substantively modify a person's attitudes toward false assertions. That is, one may not judge a false assertion to be true simply because one judges one or more of its conversational implicatures to be true. For how, DeRose asks, could a true implicature substantively modify our attitude toward a false assertion? For, except where we engage in special practices of misdirection, like irony or hyperbole, don't we want to avoid falsehood both in what we implicate and (especially!) in what we actually say?

Like the first constraint, however, this second constraint also fails to disqualify the invariantist's appeals to knowledge WAMs as a resource with which to avoid the need for contextualist analysis, even if this is for reasons less drastic than the threat of argumentative circularity. Jessica Brown notes in detail that it is a strategic commonplace in the philosophy of language to maintain that false utterances can seem true as a result of conveying pragmatically true consequences. (Jessica Brown, "Contextualism and Warranted Assertibility Manoeuvres," *Philosophical Studies* 130, 3 (1996)): 407 – 435). Along more prosaic lines, it is easy to articulate

ordinary examples of how a person's rush to judgment leads one to judge a false proposition true because the situation at hand turns a person's focus to a true conversational implicature.

The third constraint, clearly presupposed by the first and second above, is that WAM-amenable assertions only seem true (or false) because they conversationally imply other propositions that actually are true (or false). This constraint, however, gives DeRose no critical pause. Unger's skeptical invariantism is easily supplemented with a WAM-type theory of mechanism, on which "know" is a term which conversational conventions allow us to positively employ even though their hyper-stringent conditions of application are seldom, if ever, met. For example, we may falsely judge speakers to know propositions only because of said propositions' true conversational implicatures concerning warranted belief and action. The fourth constraint, again, is that such conversational implicatures occur in accordance with general and systematic (rather particular and *ad hoc*) rules of conversational implicature. DeRose warns against the use of "bare" WAMs, which purport to explain away problematic intuitions of truth or falsehood without further explaining why true assertions are unwarranted, or false ones are warranted. At first sight, DeRose suggests, this demand for general explanation may seem to be met by Unger's account. The crux of Unger's account, after all, is an Absolute Term Rule (ATR), according to which "know" is a member of a wide class of terms (e.g., "flat," "straight," "empty"), which conversational conventions allow us to positively employ even though their hyper-stringent conditions of application are seldom, if ever, met. But this assessment, DeRose maintains, is premature. For, even though ATR applies "to a very wide stretch of ordinary language," it does not satisfy the generality constraint as thoroughly as do the conversational implicature rules governing other sentences to which WAM strategies are more clearly appropriate. In particular, it stands in contrast to Grice's "Assert the Stronger" rule, which dictates that, prima facie, speakers should assert the stronger of two claims when they are in a position to assert both. In contrast to the absolute term principle, this rule is language-wide in its application and clearly supported by numerous non-problematic cases. The consequence of this failure, DeRose maintains, is a significant motivational blow to absolute term principle-based invariantist accounts. For, "by not utilizing a thoroughly general rule which has clearly correct applications [like the "Assert the Stronger" principle] the Unger of lgnorance loses a lot of leverage in advocating his view." (See DeRose, *The Case for Contextualism*, 125). DeRose's criticism here is that ATR-based invariantism suffers for lack of a fully general theoretical justification. Perhaps its error-theoretic implications would be acceptable if they were inevitable concomitants of some fundamental and language-wide maxim of conversational implicature. The "Assert the Stronger" rule is such a principle, applying far and wide across all the statements of our language. But the Absolute Term rule has no such authority. Although applying across a wide selection of absolute terms, it presumably fails to reflect any central facts concerning the fundamental point and purpose of conversational exchange. Thus, concludes DeRose, "it's difficult to see where the pressure to accept a demanding invariantist account will come from," since a general contextualist account of allegedly 'absolute' terms is available which avoids systematic falsehood" (See Derose, *The Case for Contextualism*, 125). Contextualism avoids error theory; and in so doing, it proves itself to be the more elegant and intuitive option.

We would do well, however, to question whether the generality of ATR really does suffer in comparison to that of the "Assert the Stronger" maxim? I suggest that it does not, and that this is clear once we look critically at the latter principle. There are two things that DeRose's could mean in claiming that the "Assert the Stronger" rule is more general than the Absolute Term Rule.

The first is that the "Assert the Stronger" rule potentially applies to vastly more declarative sentences of natural language than does the AT rule, since the latter applies only to sentences containing "absolute" terms. The second is that the "Assert the Stronger" rule proves more inviolate in the specific cases where it is pertinent than the AT rule proves to be in the specific cases where it is pertinent. On this second reading, the former rule is more "general" than the latter rule in the sense of being more likely to trump other rules of implicature that may also apply. This would make the former rule's imperative to assert the stronger of available claims less susceptible (than the latter's imperative to use absolute terms in pragmatically approximate ways) to being over-ridden by conflicting applicable rules of conversational implicature and use.

Of these two readings of DeRose's claim that the "Assert the Stronger" rule is the more general principle, however, neither is particularly useful to his purposes. The first is of little use for being beside the point. The relevant sense in which the generality of a principle of conversational constraint might render it an effective player in warranted assertability arguments must surely regard its propensity to win out over alternative rules, not its mere availability to being raised for consideration. The second reading of DeRose's comparative claim is of little use to DeRose simply because it is implausible. This is because numerous conversational situations regularly obtain in which the very last thing a speaker is expected to do is abide by the "Assert the Stronger" rule. The situations at issue here are not restricted to those created by the usual suspects (e.g., rhetorical devices such as irony and sarcasm) (See Grice, *Logic and Conversation*, 34). These cases appear often enough to merit counter-examples status. However, we do not need to focus on instances in which understatement functions quite so closely to the surface of our talk. More interesting cases include bureaucratic and diplomatic conversational exchange. Think of pronouncements by policy Czars and administrators, such as American Federal Reserve Board chairman Alan Greenspan, who once famously uttered, "The developing protectionism regarding trade and our reluctance to place fiscal policy on a more sustainable path are threatening what may well be our most valued policy asset: the increased flexibility of our economy, which has fostered our extraordinary resilience to shocks." (Alan Greenspan, “Opening Remarks," in *The Greenspan Era: Lessons for the Future* (Federal Reserve Bank of Kansas City, Kansas City, Mo., 2005), 8) Such utterances are crafted to violate the "Assert the Stronger Rule" in an intentional, indeed a practiced, fashion. Understatement and lack of specificity are here willfully employed as impediments to univocal interpretation. This is to keep such utterances from unduly influencing the phenomena they concern, as when one does not want a person’s predicted policy actions to be effectively priced into the economy. Alternatively, think of instances of diplomatic exchange, which also often stand in direct reproach toward any supposed language-wide imperative to assert the stronger of available claims. In these cases, yet again, the effectiveness of an exchange is likely to be no less subordinate to other imperatives than to the "Assert the Stronger" rule, e.g., the imperatives to display silence, reticence, respect for elders, and to generally leave egos unharmed. More intricately, diplomatic proclamations may be crafted, not to the end of providing maximum possible information by "asserting the stronger," but to the end of understating known facts and decisions in order to gauge reactions and vet possibilities in advance.

These complications should at least make us wary of DeRose's presupposition that the "Assert the Stronger" rule serves as a shining exemplar of a "fully general" rule of conversational implicature, against which the AT rule must invariably suffer by comparison. To this extent, DeRose has not provided a convincing case for his contention that knowledge claims are not

WAM-amenable by virtue of some distinctive failure to satisfy his fourth GI constraint.

In this endnote, I have argued that DeRose gives us no convincing reason to conclude that knowledge claims fail to be WAM amenable. In consequence, he gives us no convincing reasons to think that we cannot use a warranted assertability account to explain away, rather than assume the truth of, our ordinary seemingly non-invariantist practices of knowledge attribution.

DISCUSSION NOTES/ DEBATE

CONJUNCTION CLOSURE WITHOUT FACTIVITY: REASSESSING THE HYBRID PARADOX

Jakob KOSCHOLKE

ABSTRACT: Francesco Praolini has recently put pressure on the view that justified believability is closed under conjunction introduction. Based on what he calls 'the hybrid paradox,' he argues that accepting the principle of conjunction closure for justified believability, quite surprisingly, entails that one must also accept the principle of factivity for justified believability, i.e. that there are no propositions that are justifiably believable and false at the same time. But proponents of conjunction closure can do without factivity, as I argue in this short note. A less demanding principle is available.

KEYWORDS: justified believability, conjunction closure, factivity, lottery paradox, preface paradox, hybrid paradox

It is a well-known fact among epistemologists that the following three individually plausible principles for justified believability, when taken together, give rise to paradoxes, the most popular being Kyburg's lottery paradox or Makinson's preface paradox:[1]

> Sufficiency. For any epistemic agent A, if a proposition p is very probable given A's evidence, then A is justified to believe p.
>
> Conjunction Closure. For any epistemic agent A and any two propositions p and q, if A is justified to believe p at time t and A is justified to believe q at t, then A is also justified to believe their conjunction $p \,\&\, q$ at t.
>
> No Contradictions. For any epistemic agent A, A is never justified to believe a logical contradiction, i.e. a proposition of the form $p \,\&\, \neg p$.

Recently, however, Francesco Praolini has argued that already *two* of these principles, namely Conjunction Closure and No Contradictions, lead to what he calls 'the hybrid paradox,' a new paradox sharing features of the lottery and the preface. Here is the set-up:

[1] See Henry E. Kyburg, *Probability and the Logic of Rational Belief* (Middletown: Wesleyan University Press 1961) and D. C. Makinson, "The Paradox of the Preface," *Analysis* 25 (1965): 205–207.

> Imagine that you have just completed a book that contains sentences that express all and only logically independent propositions that you are justified to believe. Because of that, *ex hypothesi*, for each sentence s_i in the body of the book, you are justified to believe that s_i *is true*. [...] Imagine, further, that you have submitted your manuscript to Perfectly Omniscient Press, and that its perfectly omniscient referee has reviewed it. Imagine that, following the policy of Perfectly Omniscient Press, the perfectly omniscient referee writes in his report that there is exactly one mistake in the book, without telling you, however, which claim is false. Assuming that you know that the referee of Perfectly Omniscient Press is perfectly omniscient, as soon as you read the referee report you come to know—and thereby justifiably believe—that there is exactly one mistake in the book. Given that you know—and justifiably believe—that there is exactly one mistake in the book, you are justified to believe that it is not the case that s_1 is true and s_2 is true ... and s_{n-1} is true and s_n is true.[2]

For brevity, let J(p) state that p is justifiably believable for me. It then holds by assumption:

(1) J(s_1) & ... & J(s_n)

And iterated application of Conjunction Closure yields:

(2) J(s_1 & ... & s_n)

But by the referee report, it also seems to hold that:

(3) J(¬[s_1 & ... & s_n])

And applying Conjunction Closure to (2) and (3) yields:

(4) J([s_1 & ... & s_n] & ¬[s_1 & ... & s_n])

Which violates No Contradictions. So, in the situation Praolini describes, Conjunction Closure and No Contradictions cannot be true together. Accordingly, to solve the paradox, we must either give up Conjunction Closure, No Contradictions or deny that the situation Praolini describes can possibly arise.

Since only few philosophers are willing to give up No Contradictions, Praolini argues that the most plausible strategy for proponents of Conjunction Closure to deny that the paradox can possibly arise is to reject (3) based on (1) and the following well-known, but quite demanding principle for justified believability:[3]

---

[2] Francesco Praolini, "No Justificatory Closure without Truth," *Australasian Journal of Philosophy* 97 (2019): 720. For another recent paradox with a similar structure see the paradox of the pill due to Marvin Backes, "A Bitter Pill for Closure," *Synthese* 196 (2019): 3773–3787) or the examples discussed in Clayton Littlejohn and Julien Dutant, "Justification, Knowledge, and Normality," *Philosophical Studies* 177 (2019): 1593–1609.

[3] A well-known exception is Priest, who *would* be willing to give up No Contradictions, see

> Factivity. For any epistemic agent A and any proposition p, if A is justified to believe that p, then p is true.

How does this strategy work? Praolini explains it as follows:

> if Factivity is true, then no one can be justified to believe that the book contains a mistake. It is easy to understand why. Remember that the paradox discussed in this section asks us to imagine that you have written a book containing sentences that express all and only logically independent propositions that you are justified to believe. Because of this, *ex hypothesi*, you are justified to believe, of each of the claims s_i in your book, that s_i is true. Then, assuming that justified believability is factive, if you have justification for the truth of s_i, s_i must be true. Therefore, none of the claims in the book can be mistaken. For this reason, it also follows from Factivity that it is impossible to be justified to believe that the book contains a mistake.[4]

More formally, Factivity and (1) yield:

(5) $s_1 \,\&\, \ldots \,\&\, s_n$

Or equivalently:

(6) $\neg\neg(s_1 \,\&\, \ldots \,\&\, s_n)$

Then, by Factivity and *modus tollens*, we obtain:

(7) $\neg J(\neg[s_1 \,\&\, \ldots \,\&\, s_n])$

Which is the negation of (3). So, (1) and Factivity jointly refute (3).[5] Accordingly, Praolini concludes that "the paradox shows that the acceptance of *Conjunction Closure* entails the acceptance of *Factivity*."[6]

But Praolini's conclusion is unnecessarily strong, if not false. For notice that his strategy only works because Factivity logically entails (but is not entailed by) the following principle which is already *sufficient* for the refutation of (3) based on (1) and which, presumably, proponents of Conjunction Closure will happily embrace:

> Negation**.** For any epistemic agent A and propositions p_1 to p_n, if A is justified to believe that p_1, A is justified to believe that p_2, etc. and A is justified to believe that

---

Graham Priest, "What Is So Bad about Contradictions?," *The Journal of Philosophy* 95 (1998): 410–26. It is worth mentioning that Praolini also considers but quickly dismisses other potential strategies for escaping the paradox. One of them is appealing to what Smith discusses as *Principle of Differential Defeat* in Martin Smith, "The Hardest Paradox for Closure," *Erkenntnis* (2020), https://doi.org/10.1007/s10670-020-00287-4.

[4] Praolini, "No Justificatory," 724.

[5] Refutation is defined as usual in terms of logical consequence: p refutes q if and only if p logically entails $\neg q$.

[6] Praolini, "No Justificatory," 724.

p_n, then A is not justified to believe $\neg(p_1 \,\&\, \ldots \,\&\, p_n)$.[7]

To see that Negation is in fact entailed by Factivity, assume that the latter holds while the former is false, i.e. assume that there are p_1 to p_n such that:

(8) $J(p_1) \,\&\, \ldots \,\&\, J(p_n) \,\&\, J(\neg[p_1 \,\&\, \ldots \,\&\, p_n])$

Then, applying Factivity to each conjunct, we obtain:

(9) $(p_1 \,\&\, \ldots \,\&\, p_n) \,\&\, \neg(p_1 \,\&\, \ldots \,\&\, p_n)$

Which is a logical contradiction. It is also easy to see that Negation is sufficient to refute (3) based on (1): simply apply Negation to (1) and the negation of (3) follows.

But if embracing Negation is enough to stop the hybrid paradox from arising, then there is no need for proponents of Conjunction Closure to embrace a principle as demanding as Factivity. They can simply embrace Negation instead—in fact, they *should*, if they also embrace No Contradictions, for Conjunction Closure and No Contradictions jointly *entail* Negation. To see this, assume that Conjunction Closure and No Contradictions are true while Negation is false, i.e. assume that there are p_1 to p_n such that:

(10) $J(p_1) \,\&\, \ldots \,\&\, J(p_n) \,\&\, J(\neg[p_1 \,\&\, \ldots \,\&\, p_n])$

Then, by multiple applications of Conjunction Closure, we get:

(11) $J([p_1 \,\&\, \ldots \,\&\, p_n] \,\&\, \neg[p_1 \,\&\, \ldots \,\&\, p_n])$

Which obviously violates No Contradictions. But if proponents of Conjunction Closure have a less demanding alternative to Factivity, then Praolini's claim that "the acceptance of Conjunction Closure surprisingly implies the acceptance of the thesis that justified believability is factive" is not true.[8]

There is, however, *some* truth in Praolini's claim. For Conjunction Closure does, together with the widely-accepted No Contradictions and a *further* principle for justified believability that might serve as a replacement for Sufficiency, entail

[7] Negation can be seen as a generalization of principle D_J which figures in Rosenkranz's structural account of justification, see Sven Rosenkranz, "The Structure of Justification," *Mind* 127 (2018): 629–629. Loosely speaking, it states that if some proposition p is justifiably believable, then its negation $\neg p$ is not.

[8] Praolini, "No Justificatory," 716. An anonymous referee raised the worry that embracing Negation instead of Factivity might lead to what Praolini calls 'maximally radical skepticism,' i.e. the view that one is not justified in believing *any* proposition. Praolini suggests that this is the case: "all other viable explanations imply radical scepticism" (724). However, I do not see how this would follow. After all, the reasoning presented here, just like Praolini's, starts with (1) as a premise in order to refute (3). Accordingly, the set of justifiably believable propositions is assumed to be *non-empty*.

Factivity. This principle can be considered a truth norm for justified believability and is obviously the converse of Factivity:[9]

> Truth. For any epistemic agent A and any proposition p, if p is true, then A is justified to believe that p.

This principle might be appealing to both proponents and opponents of Sufficiency. Proponents might find it attractive because its basic idea is closely related to Sufficiency: if for them, being *very likely* true is already sufficient for justified believability, then *being* true should be sufficient on their view, too. Opponents might also be sympathetic to Truth if their reason for rejecting Sufficiency is that its antecedent is too weak and accordingly, that the standard for justified believability is too low. On their view, something stronger than high probability is required for justified believability. And this something could be the truth of the proposition in question.

Now, to see that No Contradictions, Conjunction Closure and Truth jointly entail Factivity, assume that the three former are true while the latter is false, i.e. assume that for some p it holds that:

(12) $J(p) \,\&\, \neg p$

Applying Truth to the second conjunct, we get:

(13) $J(p) \,\&\, J(\neg p)$

And by Conjunction Closure:

(14) $J(p \,\&\, \neg p)$

Which contradicts No Contradictions. Hence, No Contradictions, Conjunction Closure and Truth jointly entail Factivity.[10]

Time to summarize. Praolini has drawn our attention to an interesting new potential paradox for justified believability. But the conclusion he draws from it is unduly strong. There is nothing that forces proponents of Conjunction Closure to accept a principle as demanding as Factivity. In fact, a less demanding principle is available. And this principle should be very attractive to proponents of Conjunction

[9] For instance, Boghossian discusses a version of this norm where justified believability is understood as epistemic permissibility, see Paul A. Boghossian, "The Normativity of Content," *Philosophical Issues* 13 (2003): 31–45.

[10] This obviously entails that justified believability collapses to truth. Notice that there is more that can be said about the interconnections between the principles discussed in this note. For instance, Factivity not only entails Negation but also No Contradictions, and together with Truth it entails Conjunction Closure. Such details are, however, left for future research.

Closure. Still, if proponents of Conjunction Closure also accept No Contradictions and Truth, then the acceptance of Factivity follows.[11]

[11] I would like to thank two anonymous referees and Roman Heil for a number of helpful comments and a vivid discussion. This work was funded by the University of Hamburg as part of the Excellence Strategy and by the German Research Foundation (DFG) as part of grant SCHU 3080/3-1 to Moritz Schulz.

NOTES ON THE CONTRIBUTORS

Javier Anta is a Student Member of LOGOS Research Group and BIAP (Barcelona Institute of Analytic Philosophy) who is currently working on his PhD thesis "Conceptual Foundations of Landauer's Principle and the Physics of Information" under the supervision of Carl Hoefer. This project is funded by a four-year FPU grant (Spanish Ministry of Education). Currently, his main area of philosophical interest lies in the field so-called 'Physics of Information,' both in its conceptual foundations (which includes its ontic and epistemic implications) and in its historical roots. He has published other recent work in philosophy of thermal physics "Can informational thermal physics explain the approach to equilibrium?" (*Synthese*, 2021) and "The epistemic schism of statistical mechanics" (*Theoria*, forthcoming). Contact: antajav@gmail.com.

J. Spencer Atkins is currently at work on a PhD at Binghamton University in Binghamton, New York. He received his M.A. from the University of Tennessee at Knoxville. Spencer's research interests include epistemology, social epistemology, ethics (especially applied and environmental ethics), and the philosophy of religion. He has articles published articles in *Environmental Ethics* and *Episteme*. Website: jspenceratkins.weebly.com. Contact: jatkins4@binghamton.edu.

Hamed Bikaraan-Behesht is a postdoctoral researcher at Iranian Institute of Philosophy (IRIP) where he received his Ph.D. in philosophy of science. In 2017, he visited Australian National University. He will join National Research Institute for Science Policy (NRISP). He works mainly on issues related to the relation between philosophy and science in areas of meta-philosophy, philosophy of science, philosophy of mind, and ethics of science and technology. His main interests lie in naturalism, physicalism, evolutionary approaches in philosophy, and the role of values in science. Contact: h.bikaraan@irip.ac.ir; h.bikaraanbehesht@yahoo.com.

Jakob Koscholke is a postdoctoral researcher at the University of Hamburg where his project on the value of knowledge is funded as part of the Excellence Strategy. Before that, he was a member of the DFG Emmy-Noether Group Knowledge and Decision and the DFG Project Probabilistic Models of Coherence and Positive Relevance. His papers have appeared in *Analysis*, *Mind*, *Philosophical Studies* and *The British Journal for the Philosophy of Science*. His joint book with Michael

Schippers on probabilistic models of coherence has recently been published by de Gruyter. Contact: jakob.koscholke@uni-hamburg.de.

Andreas Stephens is currently studying for a PhD in theoretical philosophy at Lund University. His research is focused on the intersection of naturalistic epistemology and cognitive science. Contact: andreas.stephens@fil.lu.se.

Ron Wilburn received his B.A. in philosophy from the University of California, Los Angeles, and his Ph.D. in philosophy from the University of Pennsylvania. Now retired from the University of Nevada, Las Vegas, he has published numerous papers, mainly concerning epistemological and realist issues, with an emphasis on skepticism. Contact: wilburn.ron@gmail.com.

LOGOS & EPISTEME: AIMS & SCOPE

Logos & Episteme is a quarterly open-access international journal of epistemology that appears at the end of March, June, September, and December. Its fundamental mission is to support philosophical research on human knowledge in all its aspects, forms, types, dimensions or practices.

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

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

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

Logos & Episteme is published and financed by the "Gheorghe Zane" Institute for Economic and Social Research of The Romanian Academy, Iasi Branch. The publication is free of any fees or charges.

For further information, please see the Notes to Contributors.

Contact: logosandepisteme@yahoo.com.

NOTES TO CONTRIBUTORS

1. Accepted Submissions

The journal accepts for publication articles, discussion notes and book reviews.

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

There are no submission fees or page charges for our journal.

2. Publication Ethics

The journal accepts for publication papers submitted exclusively to *Logos & Episteme* and not published, in whole or substantial part, elsewhere. The submitted papers should be the author's own work. All (and only) persons who have a reasonable claim to authorship must be named as co-authors.

The papers suspected of plagiarism, self-plagiarism, redundant publications, unwarranted ('honorary') authorship, unwarranted citations, omitting relevant citations, citing sources that were not read, participation in citation groups (and/or other forms of scholarly misconduct) or the papers containing racist and sexist (or any other kind of offensive, abusive, defamatory, obscene or fraudulent) opinions will be rejected. The authors will be informed about the reasons of the rejection. The editors of *Logos & Episteme* reserve the right to take any other legitimate sanctions against the authors proven of scholarly misconduct (such as refusing all future submissions belonging to these authors).

3. Paper Size

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

4. Manuscript Format

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

5. Official Languages

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

6. Abstract

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

7. Author's CV

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

8. Review Process

The reason for these requests is that all articles which pass the editorial review, with the exception of articles from the invited contributors, will be subject to a strict double anonymous-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 submissions will be sent to at least two reviewers recognized as specialists in their topics. The editors will take the necessary measures to assure that no conflict of interest is involved in the review process.

The review process is intended to be as quick as possible and to take no more than three months. Authors not receiving any answer during the mentioned period are kindly asked to get in contact with the editors.

The authors will be notified by the editors via e-mail about the acceptance or rejection of their papers.

The editors reserve their right to ask the authors to revise their papers and the right to require reformatting of accepted manuscripts if they do not meet the norms of the journal.

9. Acceptance of the Papers

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

10. Responsibilities

Authors bear full responsibility for the contents of their own contributions. The opinions expressed in the texts published do not necessarily express the views of the editors. It is the responsibility of the author to obtain written permission for quotations from unpublished material, or for all quotations that exceed the limits provided in the copyright regulations.

11. Checking Proofs

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

12. Reviews

Authors who wish to have their books reviewed in the journal should send them at the following address: Institutul de Cercetări Economice şi Sociale „Gh. Zane" Academia Română, Filiala Iaşi, Str. Teodor Codrescu, Nr. 2, 700481, Iaşi, România.

The authors of the books are asked to give a valid e-mail address where they will be notified concerning the publishing of a review of their book in our journal. The editors do not guarantee that all the books sent will be reviewed in the journal. The books sent for reviews will not be returned.

13. Copyright & Publishing Rights

No manuscripts will be returned to their authors. The journal does not pay royalties.

14. Electronic Archives

The journal is archived on the Romanian Academy, Iasi Branch web page. The electronic archives of *Logos & Episteme are* also freely available on Philosophy Documentation Center web page.