volume III ♦ issue 3

2012

Logos & Episteme

an international journal
of epistemology

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

INSTITUTUL EUROPEAN

Contact address:

Institutul de Cercetări
Economice şi Sociale „Gh.Zane”

Iaşi, str.T.Codrescu, nr.2, cod 700481
Tel/Fax: 004 0332 408922
Email: logosandepisteme@yahoo.com

www.logos-and-episteme.proiectsbc.ro

Logos & Episteme

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

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

TABLE OF CONTENTS

Articles

Debate

ARTICLES

BEYOND MODES OF OBJECTIVITY

Robert ALBIN

ABSTRACT: Frege, and others who followed him, stressed the role of fallibility as a means to defining 'objectivity.' By defining objective judgments as fallible, these philosophers contributed to the consolidation of a theory of objectivity which suggested interpreting epistemological, as well as other judgements, as being objective. An important philosophical implication of this theory lies in its disclosure of the inter-relations between truth and objectivity. In light of this insight, and based on an analysis of instances of false (epistemological and other) judgments, I show that truth and objectivity go hand-in-hand, while falsity and objectivity do not. This finding alone indicates the necessity to revise the theory of objectivity.

KEYWORDS: objectivity, subjectivity, epistemology, fallibility, truth

Thales is recognized as the first Western philosopher because of his claim that discrepancies exist between one's own and others' perceptions of things, and what, in reality, things are. In claiming that "all things are water," Thales was in effect saying that while the world appears to be made up of many different types of things, in effect there is only one thing; for, in the end, everything is, in reality, made up of water. In this claim, Thales drew a distinction between appearance and reality, a distinction that was to become the basis of Western philosophy and science: 'reality' is not simply what seems to be true, but that which is objectively true.

My aim here is to explore the extent to which objectivity and truth are inter-related notions. I defend an account of objectivity which shows it to be a notion whose meaning is context-dependent rather than constant. This account of objectivity does not imply that any claim can be simultaneously regarded as being both objective and subjective. Rather, it means that what is objective in one context is not necessarily so in another. I shall also argue against a notion that has been defended by Robert Nozick and Amartya Sen, that false claims can, in many contexts, be regarded as objective. In light of this insight, and based on an analysis of instances of false (epistemological and other) judgments, I show that truth and objectivity go hand-in-hand, while falsity and objectivity do not. This finding alone indicates the necessity to revise the theory of objectivity.

I distinguish between two types of contexts in referring to 'objectivity': the epistemic context, in which we evaluate claims about the world (in philosophy and the sciences, for example), and the judgmental daily context, in which we

make normative judgments about artistic, moral, professional, and legal issues, as well as about peoples' character and the worthwhileness of engaging or not engaging in certain acts.

I. The Epistemic Context

The notion of objectivity is often used in a context which presupposes the existence of a world which is independent of our perception of it.[1] This notion of objectivity is not a single one but rather includes multiple versions. Frege, like Plato, believed in the inherent objectivity of abstract entities (ideas). Others, like Aristotle, considered objectivity in more concrete terms; they conceived of the world as consisting of everyday entities, the existence of which is independent of our perception of them. Locke added to this discourse the concept of 'primary' qualities. In Kant's view of objectivity, the world of phenomena is governed by inviolable rules that apply equally to all perceivers. The linguistic turn in philosophy helped mitigate some of the conceptual problems caused by ontological imperatives and epistemological considerations in formulating philosophical views. However, towards the end of the last century, various philosophers, headed by Richard Rorty in the English-speaking world, contended that the epistemological vocabulary concerning objectivity should be replaced by a normative one. Rorty went so far as to recommend replacing talk about objectivity with talk about social invariance. More specifically, he recommended replacing talk about objectivity with talk about solidarity.[2] Then again, some analytically inclined philosophers argued that objectivity should be considered as a property of descriptions.

Despite the many different attempts by philosophers to define objectivity, epistemologically or normatively, they all share one fundamental presupposition: they all presuppose objective claims to be fallible. Frege, and others who followed him, stressed the role of fallibility as a means to defining 'objectivity.' By defining objective judgments as fallible, these philosophers contributed to the consolidation of a theory of objectivity which suggested interpreting epistemological, as well as other judgements, as being objective. An important philosophical implication of this theory lies in its disclosure of the inter-relations between truth and objectivity. In a more detailed manner, Frege claimed that objectivity is the feature of being able to justify a claim according to some external criterion.[3]

[1] Thomas Nagel, *The View from Nowhere* (New York: Oxford University Press, 1986).

[2] Richard Rorty, *Contingency, Irony, and Solidarity* (Cambridge: Cambridge University Press, 1989).

[3] Gottlob Frege, *The Foundations of Arithmetic*, trans. J. L. Austin (Oxford: Basil-Blackwell, 1954)

According to this view, an objective claim can only be made if there is a criterion for establishing its truth value. The absence of such a criterion, according to this view, renders a claim subjective. For example, consider the following two claims:

1. I have a headache.
2. I think I have a headache.

Despite the linguistic difference, semantically, the meanings of the two claims are identical. The phrase 'I think' (in claim 2) plays no semantic role. A person cannot be mistaken about her or his own sensation of pain. Whatever she or he senses is true by virtue of that very sensation alone. Unlike their objective counterparts, subjective claims admit of no gap between the world and the one who experiences it. Therefore, what seems to be the case in a subjective claim actually is the case. Claim 1 is subjective, and as such, differs from a claim such as, 'Today the sky is blue,' the truth value of which is determined by comparing it with reality and not with one's sensations or feelings. Asserting this on a cloudy day would, of course, reveal its falsity. However, as I have just adumbrated, objectivity is equated with fallibility; one would be justified in regarding the claim 'The sky is blue' on a cloudy day as an objective, albeit false, claim.

Amartya Sen has argued that the truth of an objective claim can be ascertained by resorting to a particular criterion, such as to whether it corresponds to external reality.[4] In his view, a false claim can be regarded as objective, insofar as the truth value of the claim is distinguished from its objectivity mode. Sen's account thus posits two criteria for every claim, one for determining whether it is true or false and another for determining whether it is objective or subjective.

To exemplify Sen's perspective, let us take the statement that 'The Earth is stationary while the Sun revolves around it.' Five hundred years ago, this statement was universally acknowledged to be true. Today, Sen would regard that statement as false but at the same time objective. How can his position be justified? Objectivity, Sen claims, displays a sort of invariance; in his view, objectivity is not so much a 'view from nowhere,' but a 'view of no one in particular.' Observational claims, he asserts, can be both position-dependent and person-invariant.[5] In other words, to qualify as objective, a perceptual claim must be one that would be accepted by everyone who views it under given conditions of observation. Sen assumes that all people observing a particular object from a particular fixed vantage would have identical perceptions of the object observed.

[4] Amartya Sen, "Positional Objectivity," *Philosophy and Public Affairs* 22, 2 (1993): 126-145.
[5] Sen, "Positional Objectivity," 129.

In his discussion of objectivity, Nozick takes a different view of invariance, in which he relates directly to its truth value: "An objective fact is invariant under various transformations. It is this invariance that constitutes something as an objective truth…."[6] He concedes that the ability to distinguish between correct and incorrect transformations is a function of empirical knowledge, unrelated to *a priori* considerations. Yet, the most interesting point about Nozick's position is his assumption of the close relationship between truth and objectivity. Truth, for Nozick, is the hallmark of objectivity; a claim is objective if its truth value remains constant, even after transformations have been made. For example, assuming the truth of the statement 'The Earth revolves around the Sun.' This truth would not change, even after undergoing transformations, say, for example, by astronomical observations or by calculations made from Mars. Hence, the claim would prove to be an objective one.

In short, Nozick and Sen do not share the same view. While Nozick accepts the single criterion of truth value as a means of defining objectivity, Sen asserts that the criterion for truth (such as correspondence with reality) must be distinguished from the criterion for objectivity itself (such as the invariance of persons). Arguably, Nozick's view is more streamlined than Sen's, but a number of questions remain. For example, in another context, Nozick surprisingly seems to acknowledge the possibility of objective falsities: "an objective belief can turn out to be false. (So too can a justified belief.)."[7] Not only does this statement seem to contradict what was previously understood as his denial of the existence of objective falsities, it also employs the term 'belief' in an incongruous and hence ambiguous manner. In everyday talk, there is no room for 'objective belief.' Beliefs can be justified or not justified, true, false, blind, or rational, but they cannot be objective. Further examination, however, indicates that there is only a semblance of a contradiction for Nozick, as he distinguishes between two contexts of objectivity, one concerned with beliefs and another concerned with facts. In drawing this distinction he is, in effect, putting forward two different criteria for objectivity, one for each of the contexts. While the objectivity of a fact is established by its truth, another approach is required to establish the objectivity of a belief: "…[It] is reached by a certain sort of process, one that does not involve biasing or distorting factors that lead away from the truth."[8] Objective false beliefs are misguided or inaccurate but nonetheless they are objective. The formation of a

[6] Robert Nozick, *Invariances: The Structure of the Objective World* (Cambridge: Harvard University Press, 2001).

[7] Nozick, *Invariances*, 94.

[8] Nozick, *Invariances*, 94

true belief, though anchored in reliable factors, may still be distorted by contingencies. According to Nozick, those reliable factors are what grant objective status to a belief.

In my view, Nozick's 'factors' become the norms and rules that form the basis of our beliefs. We share common norms, which help mould our practices into uniform modes of behaviour. The norms, therefore, are independent of our perceptions. Consider the sentence, 'He don't understand me'; it clearly requires grammatical correction. Only in light of rules and norms can we distinguish between true and false, good and bad, and right and wrong. Rules constitute what Wittgenstein called 'language-games.' They function as construed modes of cultural practice and as such, are part and parcel of the context in which we make our epistemological judgments.[9] The truth value of a claim such as 'Water boils at 100°C' is to be determined by experiments carried out according to a concrete set of rules of verification. These rules, in turn, form the stage upon which our claims play their roles as truths or falsities. Wittgenstein viewed epistemological rules as normative in the sense that the rules of grammar are normative. Nozick acknowledges that, as reliable as they may be, norms and rules cannot guarantee the certainty of the judgments they sustain. They aspire to be part of the bedrock of rationality; yet, as Wittgenstein noted, they can be employed in a variety of ways, not all of which are correct. Nevertheless, Nozick was attracted to the notion that norms exist independently of our perceptions, and Sen would have nodded in agreement.

Both Nozick and Sen attempt to preserve the notion of objective falsities to varying extents. Regarding what Nozick called objective facts, it should be stressed that he maintained only one criterion by which the truth value, and at the same time its mode of objectivity, could be determined. The use of a single criterion for establishing both the truth value and the objectivity mode raises the question of how actually to distinguish between the two. If truth value and objectivity are indistinguishable, then 'objectivity' and 'truth' can only converge to the extent that they lose their distinct identities. This, I believe, is sufficient for dismissing the distinction between 'objectivity' and 'truth' on the grounds that it is non-productive. This is a sufficient reason for rejecting Nozick's latter view and, at same time, is also a reason calling for a re-examination of his account of false beliefs.

Prima facie, the assertion that false claims can be objective does not seem problematic. On a day-to-day level, we are not overly perturbed by objective falsities. However, upon deeper examination, the fallaciousness of objective

[9] Ludwig Wittgenstein, *On Certainty* (Oxford: Blackwell, 1998), #204, #410.

falsities emerges when we begin to question their truth value. The establishment of a claim as false renders its mode of objectivity as irrelevant, for it ceases to be of significance to ordinary people in daily life. Having established a claim as false, its objectivity mode is seen to lose its epistemological value, for there is no context in which people in the stream of ordinary life would have any use for the mode of objectivity of a false claim. Consider a train passenger mistakenly being told by another passenger that the last train will be leaving the station in five minutes. Having discovered the falsity of the claim, the passenger will have no interest in or use of its objectivity mode. Ordinarily, an interest in such a claim would focus entirely on whether it was true. This passenger, like all others, is expected to act on the basis of information consisting of true claims about timetables. If a claim is not true, its specific source is no longer of any relevance and it serves no purpose, as the only thing that really matters is whether or not a train is expected. Like Hollywood stunt designers who concern themselves with the intricate off-camera machinations unseen by viewers, philosophers are among the few concerned with such a deserted area of our epistemological day-to-day life, and only they are inclined to consider these falsities to be objective. Accordingly, and consistent with Nozick's intuition regarding the objectivity of truths, we have to admit that a false claim cannot be considered objective, and hence it is of no epistemological worth.

It could be argued that this view, which considers objective falsities as of no epistemological worth, can be refuted by biblical stories. Like many other myths, biblical stories are about events that, according to many, never took place. The claims embodied in biblical narrative are often considered objective, even though literally, they are false. Yet, even though these falsities may appear to be objective, for those who regard them as myths, they have no epistemological significance. On the other hand, their falsity does not render them insignificant to human culture. On the contrary, myths play important roles in our cultural make-up and educational endeavours. Many who take myths seriously view them as the textual core of their cultural existence, and for them, they are indeed part of the bedrock from which their cultural, psychological, and sometimes even professional identities are hewn out of. However, irrespective of the importance we ascribe to myths, we know that they portray imaginary facts and events; i.e., they are not literally true stories but false ones. That is why we label them 'myths.' They contain claims about the world, and can also be seen to reveal something about those who would relate to them. Were these claims true, they would teach us something about the world they portray. Their failure to express literally true statements undermines their epistemological value. Since mythological

descriptions are literally false, they cannot satisfy any epistemological needs, and hence their objectivity is never taken into consideration in evaluating them. We do not turn to myths to enhance our knowledge of the world.

Another example of objective falsity is the physician's diagnosis, which is sometimes mistaken. A patient may realize that something is amiss when her health condition fails to improve. She believes she suffers from kidney disease, but is diagnosed by her doctor as suffering from a liver ailment. Her condition does not improve even after receiving liver treatment, so some time later she goes to another doctor and receives a second opinion. The new diagnosis rejects the former one, and identifies her kidneys as the problem. She is put on a different type of treatment and her condition improves. Should the first diagnosis be dismissed as completely useless? Being false, it appears to be medically useless. It has no epistemological value as it teaches us nothing about our health.

Objectivity, then, plays a minor theme in the epistemological concerto. The main theme is played by truth. Our quest in this world begins with the search for truth rather than for objectivity. Objectivity can support the truth by reassuring us that it is related to the world and is not a mere individual hallucination or imaginary whim. Under these terms, truth becomes independent of any idiosyncratic perception. However, we tend to inquire into the texture of true claims to ascertain not only their truth value, but also their objectivity. Our search for objectivity is enhanced by the special place subjective claims play in our lives. Subjective claims are of no real epistemological value for those who seek knowledge about the world. Although subjective claims may be true or false, their contents are not open to any verification. A person complaining of pain may actually be feeling no pain at all. She is making a false subjective utterance. From a Wittgensteinian point of view, such a sentence clearly bears some knowledge for others, but none for the speaker. Subjective judgments are, therefore, incapable of representing the non-psychological world, but are nonetheless effective tools of self-expression and, in addition, an important means of learning about the world of others, although not through examination of their truth value. Statements such as, 'I'm happy' or 'Look at her, she's so attractive' are vehicles for expressing feelings. These utterances do not actually describe one's feelings or portray the woman mentioned. They are cultural modes of expression that we use instead of primitive shouts and gestures to respond to what we experience.

My goal so far in this paper has been to show the relative importance of objectivity and of false claims: neither has epistemological significance. Only true claims can be regarded as objective, and only under certain conditions. Only after we accept a claim as being true can we accept it as also being objective. If a

statement is not true it has no epistemological value. Only by following this line of thought can we appreciate the contribution of Nozick's work. It also provides a better interpretation of his insistence on the interrelatedness of the mode of objectivity and the truth value of a claim. As Nozick correctly pointed out, truth and objectivity are closely related, but they are not identical.

II. The Ordinary Judgmental Context

In ordinary life, people with a highly developed sense of justice or obligation to others are described as being faithful to their 'personal truth.' Such individual truth does not require reinforcement through objectivity as it merely expresses its bearer's sense of moral obligation, as opposed to her cognitive understanding of reality. This is simply another case in which 'truth' is employed in a non-epistemological context.

To take a concrete example, let us consider a woman tasting wine at a winery. She may say to herself, 'This is a great Cabernet, very good indeed!' Can such a judgment be refuted? Is there any objective possibility of justifying her judgment that the Cabernet is of good quality? In other words, is this judgment right or wrong, or is it merely a matter of personal taste, and as such, subjective? Wittgenstein reminded us to regard certain aesthetic judgments as simply utterances of enthusiasm: "[Words] such as 'lovely' are first used as interjections."[10] Referring to the wine as 'great' is consistent with this line of thought. But what about the woman's statement that 'The wine is good'? Is this an objective judgment? Searle would answer such a question with an unequivocal negation,[11] while Wittgenstein elaborates: "In learning the rules you get a more and more refined judgment. Learning the rules actually changes your judgment."[12] I find Wittgenstein's view more accurate and delicate than Searle's; applying Wittgenstein's observation to the wine tasting case, we could say that mastering the production of wine, particularly good wine, requires a great deal of professional skill and knowledge; it requires expertise in cultivating grapes, extracting their sugar, controlling the alcohol content, attaining the desired tint, etc. The standards for performing these tasks, many of which are grounded in age-old traditions, are taught in wine academies throughout the world. It is by these same standards that a wine's taste, colour, and aroma are judged. In view of these standards, the wine-tasting example demonstrates that objective judgments

[10] Ludwig Wittgenstein, *Lectures and Conversations on Aesthetics, Psychology, and Religious Belief*, ed. Cyril Barrett (Berkeley: University of California Press, 1997), I, paragraph 9.

[11] John R. Searle, *The Construction of Social Reality* (New York: Free Press, 1995).

[12] Wittgenstein, *Lectures*, I, paragraph 15.

regarding the quality of wine are possible, and as such, they can be right or wrong, and not merely true or false as in the epistemological context. In fact, our daily lives include endless usages of the categories of good and bad, in addition to the categories of true and false. At times, we say that the food in such-and-such a restaurant is good. We say that Michael Jordan was a great basketball player, if not the best to ever set foot on the court. Are these two latter judgments objective? Since they are fallible, it seems that they meet the objectivity criterion. The statement about Michael Jordan does not imply that it is impossible to disprove that he was the best basketball player ever; however, the disagreement alone does not render the statement subjective. By the same token, two mathematicians can argue about the right mathematical proof for a certain theorem. They may negate each other's opinions, but the dispute alone does not justify labelling either of their opinions as subjective. One or both of them may be wrong; but, if one of them is right, then her judgment is objective.

Michael Jordan's case seems to differ from the dispute over a mathematical proof, as arguably there is no accepted criterion for determining the best basketball player in the world, while there are accepted clear and distinctive criteria for determining the veracity of mathematical proofs. Given the absence of any such criteria for labelling Michael Jordan as the best basketball player ever, such a judgment would be subjective. Other difficulties might arise in cases where more than one criterion applies to a specific judgement, and in which the criteria may be in conflict. Multiple criteria make the justification of our particular judgments even more difficult, although not impossible. Though we may lack a shared criterion (or criteria) for determining who is the greatest basketball player, this lack does not preclude the full or partial acceptance of a criterion (or criteria) which could resolve the issue. Different groups may adopt different criteria, but once an accepted criterion is used by a community, its judgments can be regarded as objective. By analogy, consider the correct grammatical use of a native language. What is considered as a grammatical sentence in English, for example, is not necessarily considered as grammatical in Greek. Both languages have their own sets of rules for determining grammatical correctness, but this does not imply that such determinations are subjective.

The intriguing question here is whether or not a judgment that is regarded as incorrect can nonetheless be regarded as objective. It seems that incorrect normative judgments, like false epistemological judgments, do not merit the title of objectivity. For, what sense of objectivity can there be in a claim such as, 'Basketball players are allowed to take more than two steps on the court while the ball is in their hands'? This claim is a mistaken report of the rules according to

which the game is played, and not a mere subjective claim. As with false epistemological judgments, we are not interested in the objectivity of a normative claim once it is known to be incorrect. Nonetheless, in many cases people say their judgements are objective, as a means to strengthen them against being seen as subjective. In many cases, declaring a claim as being objective seems to involve no more than assigning it a barely justified superlative.

The final category of judgments that I would like to discuss, which differs from those discussed above, concerns judgments related to professional appointments. Take the example of applicants being considered for a CEO position in a big business corporation. One of the company board members favours a particular candidate for no other reason than that he likes the candidate – the board member sees the candidate as a nice and easygoing man. One may say that the board member's preference is rooted in a subjective judgment. Later, the same board member sees that he is mistaken, and he changes his mind – he sees that this man is not nice, nor is he easygoing. This example reveals that even subjective judgments are fallible, and as such, seems to undermine, if not totally negate, the fundamental criterion for distinguishing between objective and subjective claims. The blurred distinction between objective and subjective poses a dilemma: either we accept the judgment as subjective, at the expense of finding another criterion for its subjectivity, or we choose to demonstrate that it is not a subjective judgment after all.

To elaborate, I suggest distinguishing between the right judgment in selecting a candidate for the CEO position, and the manner in which a board member makes his initial judgment. A 'right judgment' should have taken into account the character traits and qualifications of the candidate, such as leadership, integrity, communication skills, and professional experience. If the board member had weighed all of these qualities against the demands of the job, he would probably have arrived at a different decision. As we well know, feelings affect our judgments and sometimes even reverse them, which is why we strive to keep our judgments on a professional plane – to overcome the problem of biased decision-making. We prefer to make our judgments according to professional rules, in this case, a set of criteria for hiring. In our example, the board member erred by choosing the wrong person for the job. Mistakes are likely in many of the judgments we make; yet, this does not justify their classification as subjective judgments. Another example of confusion between incorrect and subjective judgements would be a miscalculation by a math student. She may believe that she has arrived at the correct answer, but later realize that she had not executed the

calculation properly. This example also serves as a reminder not to confuse what is subjective with what is incorrect or false.

As is the case for other famous philosophical pairs, a significant tension is maintained between objectivity and subjectivity. Traditionally they were sought as opposites to each other. In exploring the relations between objectivity and truth (and what is correct) I wish to draw from this paper two conclusions. First, truth and objectivity go hand-in-hand, whereas falsity and objectivity do not. Objectivity is related primarily to 'truth' and 'rightness' (or 'correctness'); its supreme importance lies in the fact that its recognition requires us to use our minds in a manner that transcends the boundaries of our own consciousness and embraces the external world. This finding alone indicates the necessity to revise the theory of objectivity. Second, subjective judgements are not necessarily wrong or false judgements.

DISCUSSION ON THE CHARACTERISTICS OF ARCHAEOLOGICAL KNOWLEDGE. A ROMANIAN EXPLORATORY CASE-STUDY[1]

George BODI

ABSTRACT: As study of knowledge, epistemology attempts at identifying its necessary and sufficient conditions and defining its sources, structure and limits. From this point of view, until present, there are no applied approaches to the Romanian archaeology. Consequently, my present paper presents an attempt to explore the structural characteristics of the knowledge creation process through the analysis of the results of a series of interviews conducted on Romanian archaeologists. The interviews followed a qualitative approach built upon a semi-structured frame. Apparent data saturation was reached after four interviews within initial target group (senior researchers with institutional authority). Under these conditions a decision was made to continue the interviews within a secondary control group (young doctoral or post-doc researchers guided by members of the initial target group) in order to both verify the observed data saturation and to assess the impact of the attitude of senior researchers towards scientific research on the younger generation. The preliminary results allow to assert that Romanian archaeology is still caught in a highly conservative and intradisciplinarian manner of knowledge production with a negative effect on both new knowledge production and future specialists' education.

KEYWORDS: archaeological knowledge**,** knowledge production, Romanian archaeology

Introduction

At this point in time, the sealing of the Romanian archaeology within a descriptive attitude has brought it in the situation where it is perceived as a

[1] ACKNOWLEDGEMENT: This paper was made within *The Knowledge Based Society* Project supported by the Sectorial Operational Programme Human Resources Development (SOP HRD), financed from the European Social Fund and by the Romanian Government under the contract number POSDRU ID 56815. Its initial title was *Arheologie şi epistemologie. O relaţie imposibilă în cercetarea românească?* (*Archaeology and Epistemology. An Impossible Relationship in Romanian Research?*) which I decided to change, in order to better express the content of our work and to comply with the requirements of the present journal. Grateful thoughts are extended towards my colleagues within the project, for their willingness of helping my empiricist and positivistic mind to 'twist' the epistemological way.

cultural luxury practiced by a group of exotic intellectuals, totally oblivious to the realities of the society who offers their context of existence.

This may seem as a harsh opening statement. I believe that, in order to justify such an apparently unjust judgement, it suffices a summary look on the statistics of the European Research Council regarding the financing of fundamental research projects funded through Framework Programme 7 during 2011. Thus, from a total of 774 projects selected for funding, 142 are focused on humanistic sciences, and from these 25 belong to the SH6 panel, dedicated to the study of humanity's past. What I find interesting is that from these 25 projects 10 are dealing with archaeological topics and from these, eight have managers from Great Britain.[2]

In my opinion, the explanation for the success of Anglo-Saxon archaeology relies heavily upon one single word: theory.

This term may be quite often met in Romanian archaeological literature as well, but its meaning is ambiguous at best, varying from researcher to researcher. In most cases, Romanian archaeological theory is perceived in the manner of the German archaeological school of thought, under the influence of which it has also formed, being assimilated to methodology. In the case of Anglo-Saxon archaeology, theoretical thinking is defined by epistemology. As study of knowledge, epistemology tries to identify its necessary and sufficient conditions and to define its sources, structure and limits. From this point of view, until now and to the best of my knowledge, there have been no applications of an epistemological analysis to the Romanian archaeological practice. Since an analysis and argument on the necessity of the renovation of the archaeologists' attitude towards theoretical and philosophical introspection cannot be made in the absence of the enunciation of current perceived characteristics, I am thus compelled to reiterate them briefly.

Romanian archaeology has formed in a fundamentally positivistic context, under the influence of the German school of thought. Its current theoretical core has been formulated more than seventy years ago and still remains unchanged.[3] Numerous theoretical studies have criticized[4] this state of facts and its intellectual

[2] Statistics concerning various aspects of fundamental research funding through the Ideas programme of Framework Programme 7 are available at http://erc.europa.eu/erc-funded-projects (Accessed March 13, 2012).

[3] Ion Nestor, "Sabia de bronz de la Boiu," *Sargeția* I (1937): 155-214.

[4] Mircea Anghelinu, "Note privind teoria şi metoda arheologiei preistorice din România," in *Cercetare şi istorie într-un nou mileniu* (Galaţi: Editura Universităţii Dunărea de Jos, 2002), 36-44; Mircea Anghelinu, "De ce nu există teorie în arheologia preistorică din România," *Sargeția* XXX (2002): 39-49; Mircea Anghelinu, "Theory and Method in Romanian Prehistoric

immobility. However, this value judgement is built upon bibliographical analysis which I feel the need to verify and, hopefully, complete with empirical data.

The main topics, on which data collection through interviews will focus, will be aimed at the definition of the manner in which the process of creation, validation and transmission of knowledge is perceived by archaeologists.

As reference for the outlining of some characteristics of knowledge, such as it is perceived by Romanian archaeologists, I will be referring to the dichotomous view on new and old modes of knowledge production advanced by Gibbons et. al.[5] I will thus try to establish the context of knowledge production, its attitude towards transdisciplinarity (as defined by Gibbons et. al.), social accountability and reflexivity, and manners of quality control. I will also complete my frame of reference with the definition of the process of knowledge creation advanced by the SECI model,[6] with special emphasis on the specific characteristics of tacit and explicit knowledge.

Archaeology," in *Acts of the XIVth U.I.S.P.P. Congress, University of Liège, 2-8 September 2001*, Section 1, *Theory and Methods*, (Oxford: British Archaeological Reports- International Series, 1145, 2003), 87-93; Mircea Anghelinu, "Note privind impactul marxismului în cercetarea arheologică a preistoriei din România," *Cercetări Arheologice* XII (2003-2004): 275-304; Mircea Anghelinu, *Evoluţia gândirii teoretice în arheologia din România. Concepte şi modele aplicate în preistorie* (Târgovişte: Cetatea de Scaun, 2004); Mircea Anghelinu, "Dimensiuni naţionaliste în arheologia preistorică din România: primele decenii ale secolului XX," *Valachica* 18 (2005): 5-23; Florin Gogâltan, "'Centru' şi 'periferie'. I. Între teorie şi realitate arheologică," *Revista Bistriţei* XVIII (2004): 39-62; Florin Gogâltan, "Nevoia de teorie?" in *Centru şi periferie. Lucrările colocviului naţional, Bistriţa 23-25 aprilie 2004,* eds. C. Gaiu and H. Bodale (Cluj-Napoca, 2004), 7-16; Nona Palincaş, "On Power, Organisation and Paradigm in Romanian Archaeology before and after 1989," *Dacia* NS 50 (2006): 7-56; Nona Palincaş, "Despre conceptul de cultură arheologică şi despre gândirea normativă. Pentru o dezbatere în arheologia românească de astăzi," *Studii şi Cercetări de Istorie Veche şi Arheologie* 57, 1-4 (2006): 159-185.

[5] Michael Gibbons, Camille Limoges, Helga Nowotny, Simon Schwartzman, Peter Scott and Martin Trow, *The New Production of Knowledge. The Dynamics of Science and Research in Contemporary Societies* (London: Sage, 1994).

[6] Rodrigo Arocena and Judith Sutz, "Changing knowledge production and Latin American universities," *Research Policy* 30 (2001): 1221-1234; Femke Jansik, "The knowledge-productive corporate university," *Journal of European Industrial Training* 29, 1 (2005): 40-57; Ikujiro Nonaka and Ryoko Toyama, "The knowledge-creating theory revisited: knowledge creation as a synthesizing process," *Knowledge Management Research & Practice* 1 (2003): 2-10; Ikujiro Nonaka, Georg von Krogh, and Sven Voelpel, "Organisational Knowledge Creation Theory: Evolutionary Paths and Future Advances," *Organization Studies* 27, 8 (2006): 1179-1208; Célio A. A. Sousa and Paul H. J. Hendriks, "Connecting Knowledge to Management: The Case of Academic Research," *Organisation* 15, 6 (2008): 811-830; Jing Tian, Yoshiteru Nakamori, and Andrzej P. Wierzbicki, "Knowledge management and knowledge creation in

Method

The target group has been initially limited to experts in archaeology. Through *expert* I understood a person holding institutionalized authority, which thus possesses the capacity to influence the conditions in which archaeological research is carried out within a given institutional context. I found this limitation to be necessary due to the short and fixed time frame of my project. However, I believe that the current definition of the target group will allow us to gain maximum of information with a minimum time investment. However, since apparent data saturation appeared earlier than expected – only after four interviews, the initial target group has been extended in order to include doctoral students or young post-doctoral researchers tutored by members of the target group. I believe that the expansion of the initial target group – within which I have conducted three more interviews – will be beneficial to my research, since it will both allow to verify the initial data obtained from the target group, and supply new data regarding the manner in which the attitude of senior researchers towards the process of knowledge production influences the future generation.

Interview structure

Following the initial analysis of the target group and of the interviews' topics, I decided that the most suitable form for empirical data collection would be represented by semi-structured interviews, since through its series of open questions it allows both for freedom of expression, but it also offered the means to focus upon clearly defined subjects of discussion and to detail or clarify certain aspects through the use of probes.[7]

The structure of the interview has been built through several stages involving construction of questions, criticism of questions by peers, verification of revised question through pilot interviews, second revision of interview structure and completion with possible probes. The result of the process is presented in the table below.

academia: a study based on surveys in a Japanese research university," *Journal of Knowledge Management* 13, 2 (2009): 76-92.

[7] Bogner Alexander, Littig Beate, Menz Wolfgang, eds., *Interviewing Experts* (Basingstoke: Palgrave Macmillan, 2009); Foddy William, *Constructing questions for interviews and questionnaires* (Cambridge: Cambridge University Press, 2001); Gilham Bill, *Research Interviewing. The range of techniques* (Maidenhead: Open University Press, 2005).

Problem definition	Question	Possible probes
Definition of the archaeologists' perception on the research process	Which are the sources for the identification of new research topics?	- bibliography - fellow researchers - technological progress - social environment - political environment
	How would you define the purpose of archaeological research?	- description of archaeological material - reconstitution of ancient reality - integration of archaeological data within contemporary context
Definition of the archaeologists' interaction with his activity environment	Which are the principal means to solve the research problems?	- independent research - pluridisciplinary team research
	Which are the factors that offer and confirm the value of the archaeological research?	- peer review - new research directions - young researchers formation - real world aplicability

During interviews I continued the improvement of the data collection methodology, with an accent on the systematisation and categorical analysis of the answers.

Results

Sources for new knowledge

As summarized in the table presenting the interviews' structure, the first two questions aimed at outlining the following two main ideas: the identification of sources for new research problems and the definition of the purpose of the archaeological research.

In the first case, all seven participants in my study indicated the archaeological literature as first and most important source for the identification of new research directions. The second main source is constituted by new empirical data obtained through field research, especially archaeological excavations. An interesting problem was raised by the analysis of the role played by interpersonal interaction as a source of new research ideas. In only two of the seven interviews dialogue with other specialists has been willingly advanced as secondary source of inspiration and new ideas. In the other five cases in which I

used this notion as a probe the answers were as follows: in two cases interpersonal interaction was met with total rejection and in the other three it was accepted under the reserve of the existence of common interests within strictly delimited research topics. Within this context, the communication at an interdisciplinary level is seen as a second rank source, following intradisciplinary archaeological subjects.

Purpose of archaeological research

When asked to define the purpose of the archaeological research, the first option of all interviewed specialists has been *the reconstruction of the prehistoric life*. The detail of the idea highlighted the major attention towards economical aspects such as trade routes, subsistence strategies or technology, as well as the definition of the cultural evolution from the perspective of mutual interaction between distinct archaeological entities, an approach specific for the cultural – historical thought.

Only four of the researchers have regarded archaeology as being relevant towards a larger audience, either through the display of extraordinary finds, with a powerful visual or emotional charge, or through the elaboration of standards and policies for heritage management. None of the specialists considered that the general public might have an interest, nor have they conceived a larger spectrum of application of the extended results of the archaeological research. We must also note that the formative potential of the research activity does not constitute a concern for any of the participants in our interview.

Main methods

In the case of the methods used to acquire new data and build new knowledge, the intradisciplinarian approach is dominant. The main source for new data acquisition has been, in all cases, identified as the archaeological excavation, seconded by bibliographical research, while for data interpretation the cultural – historical perspective is dominant.

The importance of interdisciplinary investigations, although recognized as paramount, is oriented towards the verification of results obtained through conventional means. The activity of an interdisciplinary team is only seen as possible in the case where the archaeologist is capable of fully controlling the process of data acquisition and interpretation through the appropriation of the specific concepts and methodology; in one case, this idea is carried even further and the interdisciplinary research is seen as possible in the absence of a team, the

archaeologist being able to implement himself the methodology and to process the data.

Research validation

From the point of view of the seven archaeologists who participated in my study, research validation is constructed exclusively through peer-review, in its various shapes: reactions of colleagues to published results, citations, or awards and distinctions. The results of interdisciplinary research only come into play if they offer the possibility of supporting the intradisciplinary archaeological conclusions. The role of the general public has only been discussed as a result of the use of a probe and it usually was dismissed as a factor of validation of research results. I do feel compelled to note that, in the only one case where the general public has been accepted as a possible judge of the results of the archaeological research, its validity has been connected to the level of culture existent within the given society – in the case of the Romanian society, the general public's level of education has been considered insufficient in order for it to play a significant role in research validation.

Discussion

Although I have yet to complete the interpretation of the data obtained through my interviews, I consider the preliminary results to build a sufficiently clear and alarming image.

Through the answers I was offered, I must first observe that the process of knowledge production in Romanian archaeology continues to be characterized by a highly conservative attitude. The location of the sources for new knowledge within intradisciplinarian boundaries, the acceptance of only peer-review as factor for results validation/quality control and the sub-summation of interdisciplinary research to intradisciplinarian objectives, all point to the immobilisation of knowledge production in 'mode one.' This state of fact blocks the archaeological research to involve itself in a series of actions which might impact and benefit its social context. In my opinion, there are a series of fields where archaeological knowledge could bring a valuable contribution, such as: building a diachronic perspective on durable environment exploitation (with possible consequences on agricultural and forestry policies) or heritage management policies oriented towards tourism development, which are currently largely ignored. It all becomes much more evident when we try to identify the existence of a feeling of social accountability of archaeological research. Although heritage management and protection are regarded as one of the important outputs of archaeological research,

the manner in which the elaboration of their policies is perceived ignores their possible social impact and denies the right of the general public to contribute to the process or judge its outcome due to what the archaeologists perceive as the lack of societiy's education on the matter. This self sufficiency comes in flagrant contradiction to the large body of literature concerning this very sensitive topic already existent at European and global level.[8]

Furthermore, I find even more distressing the rigidity with which interpersonal communication is built. The dissemination of the various aspects of the research activity is only seen possible in explicit form through papers or communications. This attitude directly impacts on the formation of future generations; formed on the basis of values expressed through explicit knowledge, the patterns of thought of doctoral or post-doctoral researchers closely replicate those of their professor, being characterized by the linearity and uniformity specific for the still dominant cultural – historical approach within the archaeological practice.

To conclude, the results of my interviews confirm the existent criticism towards Romanian archaeological intellectual immobility and strengthen my belief in the necessity of archaeological thought reconfiguration on pragmatist foundations, which would allow a natural evolution towards 'mode 2' of knowledge production. The reorientation of archaeological research along the lines of thought of William James and Richard Rorty[9] would allow the Romanian

[8] Only a few titles for exemplification: Cristopher A. Bergman and John F. Doershuk, "Cultural Resource Management and the Bussiness of Archaeology," in *Ethical issues in archaeology*, eds. Larry J. Zimmerman, Karen D. Vitelli, and Julie Hollowell-Zimmer (Oxford: AltaMira Press, 2003), 85-97; Ian Hodder, "Archaeological Reflexivity and the 'Local' Voice," *Anthropological Quarterly* 76, 1 (2002): 55-69; Ian Hodder, *The Archaeological Process. An Introduction* (Oxford: Blackwell Publishing, 2003): 59, 63, 160-161; Cornelius Holtorf, "Paul Feyerabend: Towards a Democratic Relativism in Archaeology" with comments by Kathryn Denning and Per Cornell, in *Philosophy and Archaeological Practice. Perspectives for the 21st Century*, eds. Cornelius Holtorf and Håkan Karlsson (Göteborg: Bricoleur Press, 200), 241-259; Michael Shanks, "Archaeology/politics," in *The Blackwell Companion Guide to Archaeology,* ed. John Bintliff (Oxford: Blackwell Publishing, 2001); Laurajane Smith, *Archaeological Theory and the Politics of Cultural Heritage* (New York and London: Routledge, 2004); Michael K. Trimble and Eugene A. Marino "Archaeological Curation: An Ethical Imperative for the Twenty-First Century," in *Ethical issues in archaeology*, eds. Larry J. Zimmerman, Karen D. Vitelli, and Julie Hollowell-Zimmer (Oxford: AltaMira Press, 2003), 99-112.

[9] William James, *Essays in radical empiricism* (New York: Longmans, Green & Co., 1912), 73-74; Richard Rorty, *Philosophy and the Mirorr of Nature* (Princeton: Princeton University Press,

archaeological research to rediscover its humanity and to assume an empathic and open attitude towards the cultural dilemmas of the communities within which it finds its field of action, pushing the archaeologist towards an active role in the education of the general public and thus creating himself the premises of a wider foundation for the justification of his activity. Moreover, the acceptance of truth as a hermeneutical, continuous process of contextual dialogue, would help the archaeological enquiry break open its descriptivist corset and evolve towards a truly transdisciplinary conception of its research topics, thus increasing its capacity to incorporate within its interests some of the needs of the contemporary world.

1981); Richard Rorty, *Truth and Progress: Philosophical Papers, Vol. 3* (Cambridge: Cambridge University Press, 1998).

EVOLUTION, PSYCHOLOGY, AND CULTURE[1]

Bogdan BOGHIȚOI

ABSTRACT: My goal is to clarify the type of relations one could hope can be established between psychology and the social sciences in general, on one side, and evolutionary biology, on the other. Thus, the paper analyzes one of the most remarkable contemporary attempts to forge such ties, namely that of John Tooby and Leda Cosmides, who explore the interface between the two domains and try to articulate a research methodology aimed at their better integration. Unfortunately, as I shall try to show, the position Tooby and Cosmides advance is undermined by adaptationist assumptions they don't manage to successfully defend. In doing so, my paper picks up the threads of the current adaptationism debate and seeks to draw some of the consequences it has for psychological research. Subsequently, I will attempt to generalize the chief results of my analysis, by emphasizing a few aspects of evolutionary theory I think are key for understanding its relation with human culture. On this grounds, I will argue for a position that makes social sciences autonomous in respect to evolutionary thinking, yet preserves solid ties with evolutionary thought, securing integration with the rest of science.

KEYWORDS: intertheoretic relations, evolutionary psychology, social sciences, John Tooby, Leda Cosmides, adaptationism

Of one of the chief epistemological concerns of our time is the status of social disciplines within the wider context of science. It is a critical problem for the both sides of what has become a notoriously chronic divide. On one hand, it is a source of vexation for natural science, whose constitutive aim is to incorporate ever new territories into its domain, but encounters difficulties and even intellectual resistance when it tries to tackle whatever is touched by culture. On the other hand it is a problem for social sciences, which still have to clarify their relations with other fields of study and asses the place reserved for them in the wider edifice of human knowledge. This constitutes a persistent source of frustration for social scientists, who often see the scientific status of their disciplines questioned on various counts of disunity. Quite often, failures of social research were blamed

[1] This work was supported by the strategic grant POSDRU/89/1.5/S/62259, Project "Applied social, human and political sciences. Postdostoral training and postdoctoral fellowships in social, human and political sciences," cofinanced by the European Social Fund within the Sectorial Operational Program Human Resources Development 2007 – 2013.

by on the lack of ties with the rest of science, which deprives them of the solid grounds and the powerful principles of organization the conceptual apparatus employed the latter can provide.[2] Thus, rooting research in the principles of some discipline on the other side of divide was often seen as a way to secure for science the realm of human mind and behavior.

One of the major strategies of the last decades to bridge this gap was was to treat the human mind and behavior as shaped by natural selection, and thus attempt to ground the theories about them in one of the cornerstones of modern scientific thinking – evolutionary theory. There are notorious in this respect the efforts of sociobiology, ever since its approach entered the forefront of the intellectual debate in the mid '70s. But the appeal to Darwinian principles is much more widespread. It pops out constantly not only in endeavors such as evolutionary psychology, which although are often shy to claim their sociobiological heritage, represent clearly some of its spin-offs. Today, it is hard to find a topic in cognitive science where evolution has not been evoked and authors that have not produced, at least occasionally, their share of more or less sound Darwinian considerations. As a consequence, the clarification of the general implications of evolutionarily principles on social sciences has become a hot foundational topic in the last few decades.

No doubt, humans, with their propensities and capacities, are evolved beings. But to what extent can we exploit this idea to bridge the gap between social and natural sciences? My response comprises two parts. On one hand I will asses the hopes placed, quite commonly, on evolutionary thinking, which was more than once deemed capable to ground or guide our investigations into the mechanisms of the mind and culture. Here my argument will involve a polemic, that will run for the most part of the paper, with what is probably one of the most articulated and influential attempts to forge a set of principles and a methodology for psychology out of evolutionary considerations – namely, with the metatheoretical reflections of John Tooby and Leda Cosmides. In spite of the negative conclusions, my goal is not to isolate social sciences in an ivory tower. I am willing to accept that mental functioning and cultural phenomena are with no exception biological phenomena, and can be dealt with as a province of biology where evolutionary theory still applies. But this does not preclude the province acquiring autonomy, and abiding laws and mechanisms to be established through

[2] This is precisely the reproach voiced by the two authors I will concentrate on. See Leda Cosmides and John Tooby, "Cognitive adaptations for social exchange," *The adapted mind: Evolutionary psychology and the generation of culture*, eds. Jerome H. Barkow, Leda Cosmides, and John Tooby (New York: Oxford University Press, 1992),163-228.

a set of a self-sufficent investigations, to which the the wider science can only take notice and structure its inquiries accordingly. In this respect, at the end on the paper I shall try to build further on the results of my critique of the programme and Cosmides, and show how the social sciences integrate with evolutionary thinking.

Setting up the stage

Cosmides and Tooby reject a type of approach in the study of man, they label "the standard model of social sciences." As they present it, this model is characterized by the reliance on cultural and group practices, learned through socialization. From such a standpoint, "the features of a particular culture are the result of emergent group-level processes, whose determinants arise at the group level and whose outcome is not given specific shape or content by human biology, human nature, or any inherited psychological design. These emergent processes, operating at the sociocultural level, are the ultimate generator of the significant organization, both mental and social, that is found in human affairs."[3] Tooby and Cosmides accuse this stance of leading to stagnation, by discouraging the investigation of the "epistemological links" with the rest of science, and asserting a false idea of autonomy for their discilplines, blinding them to the role of evolution in structuring cognition.[4]

The above description seems hardly applicable to the mainstream scholars, as it was manifestly intended, to be called "standard model." We would be hard pressed to find even amongst the fiercest social constructionists one that would deny, for instance, the contribution of our organs of vision, as they happen to be shaped, to our color discriminating behavior just because the color-concepts are highly culture-dependent. Of course, scholars are not immune to preposterous presuppositions, which might sometimes inadvertently creep into their theories. But in order for the description to be adequate, the bulk of the scientists should accept at least tacitly, if not explicitly, the tenets Tooby and Cosmides decry. Yet no serious scientists would attempt to hypothesize for instance that humans might discriminate, if appropriately molded by their respective culture, into the UV spectrum, like bees. The set of presuppositions they entertain seems different for most of them, and that shows in what they consider to be meaningful inquiry. Somehow they tacitly, and, if questioned, explicitly assent to the idea that our peculiar biological being provides us with determinate capacities, and structure their research accordingly. That is why we don't see many studies into how

[3] Cosmides and Tooby, "Cognitive adaptations", 32.

[4] Cosmides and Tooby, "Cognitive adaptations", 23.

humans see in the UV spectrum. It is this body of shared tacit assumptions and methods that gives the content of scientific paradigms, more than polemical philosophical stands about what are otherwise core tennets of modern science – namely that our body conditions the mind. Of course, there are wide variations and bitter disputes about whereto nurture might extend and where nature reigns supreme, but only few would adhere to the extreme position that denies the latter all role, and that only when the authors are philosophizing and not effectively doing science. There is also no denial that quite a few would try to show that the contribution of nature has been overestimated and many or important traits we historically considered part of our biological dispositions are actually cultural artifacts. But that should not make us think that a corresponding number of scientists deny our biological makeup all role. Modern materialism at the core of normal science rejects the idea that our physical makeup is irrelevant to the mind and the ensuing behavior and, anyway, peculiar research programmes aimed at demonstrating the cultural origins of a peculiar behavioral disposition do not need to assume that all such behavior are sheer cultural products, but only some.

Whatever its historical implausibility, I am not interested in further analysing this denial of the role of human nature precisely because such a position has big troubles coping with ideas at the core of modern materialistic understanding of the world. Whether the adversary Tooby and Cosmides fight against is a straw man or not is secondary to me. I intend to keep my inquiry epistemological. From the standpoint of ideas, the view Tooby and Cosmides decry is actually a compound of two possible positions, not sufficiently distinguished. First, there is the radical denial of any role of our bodily nature in shaping the mind I have mentioned in the paragraph above. The second idea is that the study of socio-cultural factors can be carried out autonomously with respect to evolutionary biology. In this sense, one does not need to consider how, for instance, vision evolved in order to clarify how it is employed in socio-cultural contexts. That would not mean that "the inherited psychological design" is denied any causal role here, as the first position effectively holds. It means just that the question of what is evolved trait and how it evolved is inessential for for psychological or sociological inquiries, though of course we can retrace the evolutionary history of our innate capacities. Little by little, I shall show how this possibility can be given specific content.

As a successor for the "standard" approach, Tooby and Cosmides propose an "Integrated Causal Model" for social sciences, which no longer offers free hand to socio-cultural factors. According to their approach, the human mind is composed of a number information-processing mechanisms, put together through natural

selection that have the mission to solve the challenges thrown at our ancestors by the Pleistocene environmental conditions. In order to gain an insight into the inner workings of the mind, scientists should start by investigating what adaptive ends the mind serves, that is its functions, and then reverse engineer it, by establishing what structure enables it to perform those functions. This approach promotes a model of the mind which sees the psyche as a collection of punctual solutions or, more concretely, domain specific modules, put together by natural selection to meet peculiar challenges.

The method Tooby and Cosmides prescribe involves a few steps.[5] First, researchers are asked to determine, in so far as possible, what recurrent problems our ancestors faced and the informational resources they could employ to solve them. "Such features and relationships constitute the only environmental information available to whatever cognitive program evolved to solve the adaptive problem. The structure of the cognitive program must be such that it can guide behavior along adaptive paths given only the information available to it in these Pleistocene conditions."[6] This offers a set of constrains that any hypothesis about the structure of the above programs must comply with. We are thus provided with a heuristics that helps us generate hypothesis about the specific algorithms animating psychological mechanisms that must exist in order to address the problems environment threw at our ancestors. Finally, the proposed hypotheses about such the computational structure of such programs should be tested against patterns of current behavior.

In so doing, the method of evolutionary psychology brings with it a few conceptual tools it hopes it could extend the use to the mental domain. Thus, it employs a concept of function life sciences use and relates it in a specific way to what cognitive scientists name 'functional description' of a psychological mechanism. According to evolutionary biology the application of the concept of function must be reserved to the processes promoting fitness.[7] To gain in rigor, we should leave aside lay uses of the term, like those making it designate something contributing to the attainment of one's goals or to making a valid inference, for instance. Instead, we should take up the notion of function employed by

[5] John Tooby and Leda Cosmides, "Evolutionary Psychology and the Generation of Culture, Part I. Theoretical Considerations," *Ethology and Sociobiology* 10 (1989): 40-41.

[6] Tooby and Cosmides, "Evolutionary Psychology I," 41.

[7] Tooby and Cosmides, "Evolutionary Psychology I"; John Tooby and Leda Cosmides "The Psychological Foundations of Culture," in *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*, 19-136; Leda Cosmides and John Tooby "Mapping the Evolved Functional Organization of Mind and Brain," in *The Cognitive Neurosciences*, ed. Michael S. Gazzaniga (Cambridge: MIT Press, 1995),1185-1197.

evolutionary biology, where "it refers solely to how systematically caused its own propagation in ancestral environments."[8]

It is the identification of biological functions that helps individuate the psychological structures to be further analyzed in order to establish how they work. The only proper object for scientific study when it comes to functional architecture are the ones validated by this biological standard for functionality.[9] After all, "adaptive organization is the only kind of functional organization that is there to be found,"[10] as the evolutionary processes are the sole capable of coming up with complex structures. Thus "modem evolutionary biology constitutes, in effect, a foundational 'organism design theory' whose principles can be used to fit together research findings into coherent models of specific cognitive and neural mechanisms."[11]

As I already stated, I am not going to deny that nature informs our psychology and culture. For someone who accepts Darwinism, it is quite trivially true that evolution structures cognitive mechanisms and ultimately cultural behavior. What I shall attempt to show next is something different, namely that the outlook and subsequently, the method proposed by Tooby and Cosmides are a bad guide to human nature and ultimately misrepresent evolutionary theory.

Function and functionalism

I would like to start my argument by emphasizing a conceptual distinction. It is that between the term 'function,' as it is used in evolutionary biology, and what is properly called 'functional role' (sometimes too informally called 'function' of a opinion, intention and so on). They should be clearly set apart, as the notion of 'functional role' has a much broader scope. The functional role is defined with reference strictly to the causes and effects of the states fulfilling that role. But such a state, defined by its functional role, might have recurrent maladaptive effects or effects that are irrelevant to fitness. Think here of certain representations in a population, like the belief that killing witches can eradicate epidemics, expecting a concurrent clan not to attack when it is actually preparing for war, or of whatever opinion was deleterious to us or our Pleistocene ancestors.[12] Such states have a

[8] Tooby and Cosmides, "Mapping," 1187.

[9] Tooby and Cosmides, "Mapping," 1187.

[10] Tooby and Cosmides, "Mapping," 1188.

[11] Tooby and Cosmides, "Mapping," 1186.

[12] We can be pretty sure our ancestors had such beliefs. After all, all organisms have mental states that impact negatively on survival, from the fish that mistakes the bait for an insect, to

clear functional role, that is such states are caused by certain events and elicit certain behaviors. But they might have no adaptive function.

This mismatch makes the perspective professed by Tooby and Cosmides and the conceptual apparatus it employs too restrictive for science, even for evolutionary biology. One reason is that the chief interests of cognitive scientists lies precisely in the cognitive role, in the sheer causes and effects of specific psychological structures, even of those that, when evaluated from an evolutionary standpoint, must be considered as having no adaptive virtues or even being maladaptive for our ancestors. Understanding how mind works is one thing, while understanding how well it works is quite another. The roles of clearly dysfunctional states of mind even constitute some of the chief points of interest for various branches of applied and clinical research. For instance the causes and effects of the different classes of psychotic thoughts, drug produced hallucinations or injury induced dysfunctional states, for which we would hard pressed to find adaptive virtues, have always been one of the chief interests of psychiatry. For instance the phenomenon of command hallucinations, urging people to harm themselves or those around them is hardly adaptive, even for Pleistocene conditions. Of course, some of the states we currently consider as mere pathological dysfunctions could turn to be adaptations. For instance, depression might have been a selected feature.[13] As a matter of fact, nosographically isolable deviations from the regular functioning of the mind, or to be precise, pathological states possessing a specifiable set of causes and effects are bound to appear and interfere with the workings of the mind, as designed by evolution. Malfunction is something that just happens, often in a from that does not vary from individual to individual.[14] Sticking to adaptive mental processes and states would discard many respectable and valuable areas of scientific research, like mental health, robbing them of essential parts of their conceptual apparatus.

More generally, taking adaptedness as a criteria for genuine biological kinds, physical or psychological, creates at least two difficulties. First it excludes biological phenomena like diseases and recurrent malfunctions from the realm of science. Thus perfectly respectable medical statements such as "Aneurysms increase the risk of hemorrhage" would suddenly be banned from scientific publications, because an aneurysm is a structure which is deeply maladaptive, and

---

modern humans falsely believing, like the Chinese alchemists did, that ingesting mercury salts prolongs life.

[13] P.W. Andrews and J.A. Thomson Jr, "The Bright Side of Being Blue: Depression as an Adaptation for Analyzing Complex Problems," *Psychological Review* 116 (2009): 620-654.

[14] Or even presumably between species, allowing for animal models of human mental diseases

therefore, it cannot be a genuine biological structure. Secondly this stance creates problems even within evolutionary thinking. After all, it is not only physicians that are interested in maladaptive structures. Failure to adapt and biological configurations that are ill-suited for survival are part and parcel of evolutionary theorizing, without which we cannot understand natural selection. Working with such structures is required in order to explain, for instance, why certain individuals or populations died out. Therefore being an a function-performing adaptation is not a necessary condition for being a scientifically respectable biological structure, physical or psychological.

Adaptionism and antiadaptionisms: picking up the threads

But the most discussed problem any form of adaptionism faces is that of unselected features which might creep into the design of living beings. There are many factors that have been mentioned as non-selected features. For instance, there is the issue of exaptations of and spandrels. Then, there is sheer chance, which might for instance produce of fail to produce a mutation at the origin of a phenotypic trait. Tooby and Cosmides are of course well aware of such phenomena, but often they fail address the counterarguments these phenomena generate.

The spandrel problem was notoriously pushed forward by Gould and Lewontin in their *The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme*,[15] and often voiced specifically against evolutionary psychology in its later writings by Stephen J. Gould.[16] The term 'spandrel' is of an obvious architectural extraction. In its original usage, it designates a certain empty space created by adjoining an arch to the straight boundary of another architectural structure. This creates a feature unintended by the architect, who deliberately designed only the architectural structures bordering the spandrel. Traits the evolution has created for their adaptive value might have the same 'unintended,' or, to be precise, unselected features. Think for instance of the color of the bones, to take an often quoted example. They are white because of the peculiar mineral composition of the bones, which was undoubtedly selected by the evolution. But whiteness itself wasn't selected. Bone color contributes nothing to overall fitness, and as such is an unselected feature of

[15] S. J. Gould and R. C. Lewontin, "The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme," in *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 205, 1161 (1979): 581-598.

[16] Stephen Jay Gould and Elizabeth S. Vrba, "Exaptation–A Missing Term in the Science of Form," *Paleobiology* 8, 1 (1982): 4-15.

the animal organism. Similarly, many features of our psyche could be such side effects, not selected by any adaptive process.

Of course, Tooby and Cosmides are not naive to such difficulties. They address both lines of attack. First they eagerly point at the differences that, allegedly, exist between features that are due to chance and those that are the outcome of natural selection. Thus, they hold that chance is incapable of creating the complexities of human architecture. They argue that "social scientists should be extremely uneasy about positing an improbably complex structure in the system with the capacity to serve nonbiological functional ends, unless that capacity is a by-product of functionality that evolved to serve adaptive ends. Selection builds adaptive functional organization; chance almost never builds complex functional organization."[17] That is, whiteness might be a spandrel caused by the composition of the bones, but we won't get anything more complex than that.

But the argument is faulty. Of course, all complexity is the outcome of an evolutionary process. This is one of the chief theoretical conquests of Darwinism, but invoking this basic truth is of no real help here. This is because it doesn't mean that all complex arrangements were *selected*. We should make a clear distinction between what was *selected* and what *evolved*. Thus, any accumulation of features – be they, selected features, spandrels or mere accidents – would create a complex structure, that of course has an evolutionary history behind, but which might not be selected as such.

Think for instance of a human face, hand or of many other anatomical features. It is unlikely that their outer shape was selected, though some of their features of course were. For instance a certain configuration of the eyebrows is necessary for preventing sweat to trickle into the eyes. But a large part of the facial traits is the consequence of the internal skeletal, muscular or sensory structures. This means that a large number of features were not selected as such, because they are mere spandrels springing from the internal structures. We should also add to the mix whatever trait resulted from random genetic effects. Some of them are not universal, like peculiar traits that run in families due to random genetic accidents in their history. But many could have spread in an entire population through to genetic drift. Of course, science is far from sorting out which is which. Nevertheless, the occurrence of each type of process is a common event in the evolutionary history, so most likely we will have all in the mix.

This piling up of features resulting from genetic accidents, of spandrels and selected features is capable of creating complex structures. Faces and hands have a

[17] Tooby and Cosmides, "The Psychological Foundations," 110.

complex shape, that usually requires years in art school for someone to be able to reproduce in clay or draw. Thus, the argument out of complexity is not conclusive.

One might object that there is a difference between piling up various items and a coherent structure, whose internal proprieties provide it a role in the life of a species. But faces are not amorphous, in the sense that their structure is psychologically, and implicitly biologically relevant. For instance, their internal arrangement is decisive for them being recognized. The human brain even possesses well-defined structures for such a task, that for obvious reasons came to be called "the fusiform face area."[18] Thus, faces get to have an important role important consequences in the life of humans, from triggering specialized recognition mechanisms to eliciting more subtle, culturally charged reactions.

Let us remark here that the argument turns not on the possibility of randomly creating complex functional systems in the biological world. After all, given enough time such unlikely events are bound to happen. It is rather about what is a steady process in the course of evolution and what is the improbable exception. But that won't help at all the argument of Tooby and Cosmides either. First, such combinations that have gained a biological function turn out regularly. The example of the human face is not an unique in the history of evolution and similar cases pop up on many other phylogenetic branches, which shows it to be a regular phenomenon, not a chance event occurring in of one line of evolution or, at best, in a handful of them. For instance the entire body shape, sculpted by the same categories of factors, acquires a role in parent recognition for many species. Of course, there are species who have developed explicit cues for that task, such as the herring gull, which evolved a special red patch on the bill, so that it could be recognized by its chicks.[19] But the display of such selected manifest recognition cues is hardly the rule in the animal kingdom. The living world is split into two numerically consistent groups with regard to recognition strategies. Many animals do not employ such visible outer signs. For some, even any randomly shaped middle-size object would actually do. Lorentz showed that, for the youngsters of many species, the cues are provided through filial imprinting at an early age, and are not specially selected visible structures.[20] Thus, chicks will imprint any conspicuous moving object of the right size they are in contact at a critical age, regardless of the shape. They could come to take as their mother members of other

[18] J. Sergent, S. Ohta, and B. MacDonald, "Functional Neuroanatomy of Face and Object Processing. A Positron Emission Tomography Study," *Brain* 115 (1992): 15–36.

[19] N. Tinbergen and A.C. Perdeck, "On the Stimulus Situation Releasing the Begging Response in the Newly Hatched Herring Gull Chick (*Larus argentatus argentatus* Pont)," *Behaviour* 3 (1950): 1-38.

[20] Konrad Lorenz, "The Companion in the Bird's World," *Auk* 54 (1937): 245–273.

species or even inanimate objects, such as researcher's boots and even a celluloid ball. This demonstrates that for such species there is no specially evolved outer structure on which recognition would depend. Whatever form they happen to encounter at a critical age might be adopted. In the wild such a role must be assumed buy the accretion of features, of which many are unselected, that forms the outer bodily shape of their parents .

Therefore, such biological structures might come to play a biological role in many other lineages. The examples could continue indefinitely. The bodily shape of a carnivore might make pray or competing predators flee. Its scent – which is the result of a mix of selected pheromones, of components that were not selected for their odor, like disassimilation products, and whatever genetic accidents might throw in the mix – assumes a host of roles, from marking territory to signaling the presence of a possible mate. Such assemblages are recurrent in many disparate phylogenetic lines. We are not dealing here with a chance event that we could discount, as we would do with the possibility of a monkey typing a Shakespeare play, when studying primate cognition. These are forces steadily operating in all organisms, and frequently coalescing to form structures significant for many species.

There is also a second counterargument Tooby and Cosmides construct in response to the idea that many of our psychological features might be unselected. It addresses specifically the problem of spandrels. Thus, Tooby and Cosmides contend that any attempt to explain spandrel-driven behavior must ultimately resolve into an evolutionary story of how such side-effects appeared. They hold that "the explanation for any specific concomitant or spandrel consists in the identification of the adaptation or adaptations to which it is coupled, together with the reason why it is coupled."[21] For instance, the color of bones is an inevitable consequence their composition. As such, their whiteness could be given an evolutionary explanation, by pointing at the past processes that that made bones contain certain minerals and the reasons why such a composition colors them white (i.e. mentioning the optical proprieties of certain substances). Therefore, adaptionism can explain away spandrels. They no longer pose a problem to those assuming that biological structures must be shaped uniquely by natural selections, as such by-product structures can be shown once again to be the sheer outcomes of an adaptive processes.

But if we are to be precise, such explanation can be constructed only if evolutionary theory is supplemented with one or more truths extraneous it. They can be provided by other sciences, such as chemistry or physics, as are those about

[21] Tooby and Cosmides, "The Psychological Foundations," 63.

the optical proprieties of calcium compounds. Or, in order to explain the coupling, the additional statements might describe contingent facts, which allows randomness to creep in again. For instance hydrangeas have developed the capacity for the hyperaccumulation of aluminum. This ability is routinely explained as an adaptation, whenever the soil releases large quantities of that substance, in occurrence aluminum. One side effect of the high aluminum levels in the organism is that, by interfering with the chemistry of certain pigments, it alters the color of the flowers. Thus, in acidic soils, which favor aluminum absorption, the flowers will be blue. But if grown in non-acidic soil, which lowers the quantity of absorbed aluminum, the flowers turn pink or cream colored.[22] Flower color is a spandrel. It is the accompanying effect of the selected capacity to accumulate aluminum. But in order to explain the flower color of a population, or the variation of color within the members of the entire species the population belongs to, we must call in an accidental factor – the peculiar environmental conditions the plants just happens to live in, or, in the latter case, to point at the contingent, historical fact that the species managed to colonize or was cultivated in soils with a lower pH than those presumably put a selective pressure on it Only by invoking such fortuitous factors we can link plant color to evolutionary history and selection, and see how one leads to another. Color is not the effect of adaptation, not even its automatic side effect. There is more to the story than evolution. Biological features are structured by many more factors, from stable laws of nature to sheer environmental contingencies.

From environmental conditions to mental structures: adaptationism as method

Some adaptationists decided to bite the bullet and accept that non-selected components are effectively part of our psychological makeup. They came to see the idea that psychological traits are adaptations as a mere methodological presupposition we need to embrace provisionally, in order to further the evolutionary investigation of our mind. Of course, reply the partisans of this stance, not all features are selected. But if we are going to apply Darwinism, they contend, whenever we are dealing with a biological trait, we need to start by supposing that the feature under consideration is selected, so that we could construct a theory accounting for its evolution, a theory which of course, could later be falsified.[23] Or to put it as Dennett does, "it is never a mistake to ask the

[22] P. B. Larsen, "Unraveling the Mechanisms Underlying Aluminum-Dependent Root Growth Inhibition in Genes for Plant Abiotic Stress," in *Genes for Plant Abiotic Stress*, eds. Matthew A. Jenks and Andrew J. Wood (Wiley-Blackwell, 2009).

[23] Cf. John Alcock, *The Triumph of Sociobiology* (Oxford: Oxford University Press, 2009).

adaptationist's 'why' question, even when the true answer is that there is no reason."[24]

I actually have no reticence to accept that provisionally assuming adaptation, as an essential part of the construction of a larger scientific hypothesis, might yield good method whenever we know what the features characterizing a species are and, subsequently, we are seeking to establish how they appeared. But vice-versa, by assuming that there was an adaptive process in order to determine the structures of the mind, we are stepping on shaky grounds. Given our task, we shall see whether presupposing that an adaptation has taken place yields a good methodology for evolutionary psychology. Unfortunately, it doesn't, at least if we see the the method of evolutionary psychology the way Tooby and Cosmides do.

As described above, the method proposed by Tooby and Cosmides requires us to figure out ancestral environments and assume there was of an adaptive process resulting in a psychological mechanism, whose computational structure should be then deciphered. But this it is a very unreliable guide for singling out specific psychological structures and assigning them biological functions and internal architecture. Such a method for mapping the structures of the mind and brain is undermined by several factual and epistemological problems.

First of all, the picture of the human psychological evolution Tooby and Cosmides work with – one of specific environmental conditions creating specific psychological adaptations – is at least disputed if not empirically mistaken. For instance Potts[25] argues, based on a solid geological record, that the specific traits of humans took shape in an era of increased environmental variability. This determined the replacement of environment-specific adaptations with mechanisms capable to deal with the inconsistent environmental conditions. This leads him, contra Tooby and Cosmides, to assert that it "is patently incorrect to characterize the human ancestral environment as a set of specific repetitive elements, statistical regularities, or uniform problems which the cognitive mechanisms unique to humans are designed to solve. This portrait of the Pleistocene environment should be discarded and with it the view that the human mind is composed mainly of innate special purpose devices or algorithms tied to a particular array of past adaptive possibilities."[26] According to Potts, nature seems to have favored more flexible mechanisms, capable of yielding adaptive behavioral

[24] Daniel Dennett, *Darwin's Dangerous Idea* (New York: Simon & Schuster, 1995), 276.

[25] Richard Potts "Environmental Hypotheses of Hominid Evolution," *Yearbook Of Physical Anthropology* 41(1998): 93–136; Richard Potts, "Variability Selection in Hominid Evolution," *Evoutionary Anthropollogy* 7(1998): 81–96.

[26] Potts, "Environmental Hypotheses."

responses in a broader spectrum of situations. Therefore the picture of the evolution that portrays psychological mechanisms as adaptations to specific environmental conditions rests on disputed premises.

Moreover, besides the fact that the argument is based on questionable assumptions about the Pleistocene environment, the conclusions drawn by Tooby and Cosmides simply do not follow.[27] Even if we disregard the empirical data and assume that humans evolved in relatively stable environments, we cannot anticipate whether the psyche will contain adaptations specific to a peculiar class of problems. One reason is that, we cannot tell, based on environmental or comparative data whether a certain psychological adaptation has effectively evolved at all. Thus, an adaptation to a specific challenge might never appear at all, even in populations that manage to survive and reproduce. There are quite a few alternative scenarios to that of environmental pressure generating a matching adaptation. First, the lack of a specific adaptation in a certain respect might be compensated by other very successful phenotypic traits. Prairie grasses, unlike thorny shrubs or the plants that accumulate nasty tasting substances in their leaves, have no way of avoiding being eaten by grazing animals. Their success is ensured by other means, such as high fecundity and the ease of spreading, which compensates for their vulnerability. Thus, one cannot infer that certain specific adaptation has taken place – for instance avoiding being eaten by grazing animals – even when we can identify a clear selective pressure in that direction. This applies to the psychological adaptations too. For instance our ability to cooperate helped us in many tasks – for instance hunting – for which other species developed specific cognitive adaptations – like better sensory discrimination, helping them to better spot pray in dense vegetation. Though fitness might be increased by a certain adaptation, species might do without it, just because they present other overwhelmingly advantageous features. It would have been nice to spot game faster, but we managed to overcome our mediocre visual processing by organizing battues with the other members of the group, thus relying on our social

[27] Some critics of evolutionary psychology also objected that the paucity of raw data about the environments of our ancestors, precludes us from drawing any conclusions with regard to our psychological adaptations (Cf. Robert Richardson, *Evolutionary Psychology as Maladapted Psychology* (Cambridge: MIT Press, 2007)). This might be true, but it cannot be a fundamental objection. The accumulation of paleontological data might solve many of our problems here, in the same manner many other scientific enigmas in all areas of research were answered by newer and richer data. What we hold is more radical, namely that even if we knew perfectly well the environment and the populations that were the basis of selection, strictly speaking, we would be able to conclude nothing about the adaptations a certain species has accumulated.

abilities. Secondly, even where an adaptation does effectively appear, there is the often emphasized possibility that its propagation in a population might be stopped in its tracks by unfortunate events, eradicated the eruption of a volcano or epidemics wiping out the small population where the adaptive structure first makes its way into the world. Thirdly, even if we knew that a certain biological function is effectively performed, there is no guarantee that it is performed by a psychological mechanism. Take for instance protection from the harmful effects of UV. It can be obtained through a psychological mechanism that makes the individual move into the shade when insolation is at its peak, or through a purely biochemical mechanism, like secreting melanin into the skin. Again, we cannot infer the existence of a psychological mechanism. These situations do not in any way pretend to exhaust the possible range of scenarios where a certain psychological adaptation, though extremely useful, did not effectively arise. As a matter of fact, we will point to yet another important scenario at the end of our paper. What they are meant to emphasize is how many alternatives there are to a psychological adaptation, showing it to be far from the mandatory effect of a certain environmental pressure.

Unfortunately, the whole methodology Tooby and Cosmides recommend involves the assumption that there must be a program put together by evolution, whatever might be its algorithms, whose purpose is to address the problems the Pleistocene environment threw at us with the resources our ancestors had at their disposal.[28] There is nothing that guarantees us that the problem has been addressed through psychological mechanisms, to which subsequently we should unravel the inner structure and then test our hypothesis about the algorithms composing it against the patterns of contemporary behavior. Methodologically, presupposing the existence of a psychological structure for each problem the environment creates is a bad move: if we aren't guaranteed that a psychological adaptation evolved effectively, the method Tooby and Cosmides endorse will be necessarily unreliable, misleading us into following false tracks when applied consistently.

But even in the scenarios where it is true that pressure leads to some form of psychological adaptation, that fact would be of no use in guiding us through the meanders of our psyche. The method Tooby and Cosmides propose has trouble individuating the modules or structures we might posses. The reason is that often evolution chooses to go for multipurpose mechanisms. One reason is that developing an organ, be it mental or physical, for each challenge a population encounters might prove extremely costly from a metabolic standpoint. There are a

[28] Cf. Tooby and Cosmides, “Evolutionary Psychology I,” 40-41.

few viable alternatives here. One of them is a domain nonspecific structure, that solves a broader range of tasks than those required by adapting to a peculiar set of circumstances. Another option, which Gould has already emphasized,[29] consists in redeploying a structure that already exists and is employed for different tasks, but which can take up a new job, that is through exaptation. Exaptations and newly created multiple purpose mechanisms might yield mediocre results, compared to a dedicated module. Nevertheless they get the job done, enabling the organism to survive in a set environment.

Both possible evolutionary outcomes imply that we cannot hold that specifiable persistent conditions in our ancestral environment require dedicated modules. Consequently, we cannot conclude from data about a peculiar environmental pressure to the existence of a mental organ performing a determinate set of functions, and whose inner workings could presumably be further investigated by cognitive science or psychology, once isolated. In other words, we cannot identify the specific structures of the mind based upon environmental conditions. Concretely, evolutionary considerations cannot tell us how many modules or distinct structures there are, designed to cope with environmental problems, or even what is the distribution among them of the psychological functions that the mind, globally, fulfills. Once we have multipurpose devices, the whole approach Tooby and Cosmides try to construct collapses. The method Tooby and Cosmides propose is incapable to come up with a set of devices and their functions, which is a mandatory step for them to be "reverse engineered." We are in the dark to how to carve the mind into meaningful units and establish what each is for, in order to subsequently expose its inner workings, that is the algorithms they run.

Thus, the method proposed by Tooby and Cosmides fails in one of its intermediary aims, which is to individuate the structures of the mind, as a prelude for the decipherment of their computational structures. This makes the method they advance an ineffective guide to what is in our heads, as evolutionarily speaking the researcher is confronted with a variety of alternatives Applied faithfully, it would take us astray.

[29] Gould and Vrba, "Exaptation."

No royal way to the mind

So what could be the contribution of identifying the peculiar selective pressures to building hypotheses about our mind or linking psychology with the rest of science? Of course, heuristically speaking, knowing the environmental circumstances in which certain species evolved often suggests reasonable psychological hypotheses. But merely suggesting is a very unremarkable feat. The condition of suggesting scientific hypotheses can be very easily satisfied, even by the casual observation of banal everyday things. It too could suggest highly plausible or significant ideas to scientists, but that does not recommend it as a method of choice for science.

On the other hand, the capacity of evolutionary thinking to suggest psychological hypotheses makes no difference with regard to the efficacy of the thus generated research programs in forging links with the rest of science or in elucidating the mental mechanisms, compared to any other discipline that has something to do with human behavior. Things might have been different had it been shown that the generation of evolutionary psychological hypotheses were based on peculiarly strong constraints between theories. As a matter of fact, many other disciplines suggest hypotheses about the specific mechanisms of the human mind. For instance so does economics, when studying investor or consumer behavior. The kinematics of a ball thrown by a sportsman say a lot about his motor control mechanisms. But inspiring psychological hypotheses doesn't grant these disciplines, ranging from social sciences to Newtonian mechanics, a privileged position with regard to deciphering the human psyche or linking it with the rest of science. So far it looks that all these disciplines stand on equal footing, or at least no grounds for thinking otherwise were provided.[30]

There also is no denial that sometimes, due to shared cognitivist and (post)sociobiological influences or even to an explicit allegiance to a methodology, those seeking to apply evolutionary theory to mind and behavior might follow, to a certain point, the path of reasoning described by the method of Tooby and Cosmides. Generally speaking, the results that are still in wait of a definitive assessment.[31] The jury is still out to what will be the future of current attempts to

[30] Also, leaving aside suggesting hypotheses, stratigraphical or paleontological hard data and hypotheses might offer corroboration for psychological theories, but that is nothing special either. In fact any data, economic statistics, ball trajectories, PET scans, and whatever we might reasonably infer from them might do so.

[31] Tooby and Cosmides offer an example of such an application in their more applied work on the psychological mechanisms underpinning social exchange (see Leda Cosmides and John Tooby, "Evolutionary Psychology and the Generation of Culture, Part II. Case Study: A

ally evolutionary biology with disciplines dealing with mind and behavior.[32] Nevertheless, we are perfectly disposed to concede that such inquiries might get it right and come up with true theories. After all, it is as certain as it can be that our minds do contain adapted structures, and some of them might very well be single-purpose modules. Supposing that there is a module for x when such a module truly exists yields true theory. The work of many researchers is thus likely be confirmed by the future science. Other factors might contribute too to the success of their endeavors. The long familiarity of the experienced scholar with a research domain that makes her better at spotting patterns, in occurrence patterns of behavior and evolution, turns their hypotheses into more than mere shots in the dark. General scientific practices, not specific to the inquiries into behavior or evolution, applied skillfully might play their part too. But this are factors independent of the Tooby and Cosmides rulebook advance, sometimes offering guidance where otherwise there is none. They are capable to improve the efficacy of scientific inquiries, whatever the specific paths of discovery followed in a peculiar domain and whatever the domain. But applying consistently the method constructed by Tooby and Cosmides is might create confusion and, moreover, make us miss systematically what is not domain specific and is not selected.

The relations between evolutionary thinking and psychology: an assessment

Tooby and Cosmides' adaptionism is in trouble, and it cannot solve its problems even after a second round of arguments. But what is the relationship between evolutionary thinking and psychology? Strictly speaking, psychological research can dispense entirely with identifying adaptations. Evolutionary considerations will make no difference, because, as we have seen, strictly speaking they say nothing about how mind is organized and cannnot constrain the hypotheses one can advance. They imply nothing about what distinct structures the mind possess and what they're for, that is, they constrain in no way the type of hypothesis we can make.

Computational Theory of Social Exchange," *Ethology and Sociobiology* 10 (1989): 51-97), if we are to look for an example of consistent and orthodox application of their method.

[32] As a matter of fact the bold entry of Darwinism into this new arena stirred quite a bit of controversy with regard to their rigorousness of the results, regardless of the peculiar school of thought that produced them, be it the older sociobiology or the more modern evolutionary psychology; see Philip Kitcher, *Vaulting Ambition: Sociobiology and the Quest for Human Nature* (Cambridge: MIT Press, 1985) and David J. Buller, *Adapting Minds. Evolutionary Psychology and the Persistent Quest for Human Nature* (Cambridge: MIT Press, 2005).

On the other side, psychological inquiry of the regular, non-evolutionary sort is required for verifying all theory about the evolutionary processes that led to current psychological structures. This is even the final step in the methodology Tooby and Cosmides endorse, as it demands that the final hypothesis abut the mechanisms and their functioning must be checked against the current patterns of behavior (see above). Such a relation completes the conditions for an actual reversal of the roles. While evolutionary theory has nothing constraining (or peculiarly interesting compared with any other science) to say with regard to our psychological structures, the data of psychology (and of anthropology, sociology or economics for that matter) provide mandatory information about our patterns of behavior and establishing these patterns has to be done with the methods of psychology, anthropology and so on. This makes those types of investigations the ultimate arbiter with regard to the existence of a certain mechanism and the way it functions. Thus, on one hand, regular psychology and cognitive science are ultimately unconstrained by evolutionary theory, and, on the other hand, evolutionary theorizing must acknowledge its dependence of the regular psychological or cognitive theories.

Nature and culture – a wider stance

After dealing with the the detail problems of the method, I would like to turn to an ensemble view of the kind of endeavor Tooby and Cosmides advocate, as a final step towards elucidating the place of cultural behavior in the architecture of nature. From this broader standpoint, the whole adaptionist programme of Tooby and Cosmides seems to rest on the following picture of the relationship between the organisms and their environment:

> Organisms transact the business of propagation in specific environments, and the persistent characteristics of those environments determine the dangers, opportunities, and elements the organism has to use and to cope with in its process of propagation. Consequently, the structure of the environment causes corresponding adaptive organization to accumulate in the design of the organism (....) This functional organization in the organism - its set of adaptationsis designed to exploit the enduring properties of the environment in which it evolved (...) and to solve the recurring problems posed by that environment. Adaptations evolve so that they mesh with the recurring structural features of the environment in such a way that reproduction is promoted in the organism or its kin. Like a key in a lock, adaptations and particular features of the world fit together tightly, to promote functional ends.[33]

[33] Tooby and Cosmides, "The Psychological Foundations," 69.

Unfortunately, this turns out to be an one-sided view of what actually happens. Adaptation is bidirectional. It does not consist solely in a continuous shaping of the internal structures, but also in modifying the environment, making it fit the needs of the organism. As it has been already emphasized, organisms actively construct their niches.[34] For instance, by damming rivers, beavers alter the environment they will live in. Deciduous trees change soil composition to one that better suits their needs. Generally speaking, adaptation is not achieved solely trough modifications of the inner mechanisms. Sometimes adaptation is achieved by changing the external environment. Yes, the adapted populations and their environment fit like a key and lock. Nevertheless, the match may be achieved not by key cutting, but by adjusting the lock so that it would fit a preexisting key, or by adjusting both the key and the lock.[35]

This has two consequences. One is of fairly obvious import for rejecting the general picture Tooby and Cosmides embrace and the methodology it underpins. Developing cultural behaviors mitigates the problems created by peculiar environmental conditions, by altering them. There was no need to develop a dedicated psychological module that would make us survive arctic weather. We invented instead sewing and the appropriate clothing, as well as ways to build shelters, which allowed us to control the microclimate surrounding our bodies. Instead of developing mental or physical organs for finding food we created the agricultural practices that have overpopulated our environment with useful species. Of course, it is still unclear what was the range of techniques our ancestors mastered, but the point is that they did use tools. We, as well as our Pleistocene ancestors, created handy environmental items and the associated procedures to manipulate them, thus managing to reduce the environmental pressures, which might have otherwise led to psychological and physiological adaptations. Internal changes can be replaced by a modification of the environment, once the appropriate technique is devised.[36] Whenever there is

[34] See Richard Lewontin, "Adaptation," *Scientific American* 239 (1978): 156-69 and also his "Gene, Organism, and Environment," in *Evolution. From Molecules to Men*, ed. D.S. Bendall (Cambridge: Cambridge University Press, 1983), 273–285.

[35] It is at least plausible that some of this 'niche construction' might have fed back into our genetics, as some argue, leading to a process of gene-culture coevolution; see F. J. Odling-Smee, K.N. Laland, and M.W. Feldman, "Niche Construction and Gene-Culture Co-Evolution: An Evolutionary Basis for the Human Sciences," in *Perspectives in Ethology* Vol. 13, eds. Peter H. Klopfer and Nicholas S. Thompson (New York: Plenum, 2000), 89–111.

[36] Of course, tool use might have induced psychological adaptations. But the important thing here is that this is not always the case. For instance metalworking or agriculture is unlikely to have fed back into the mechanisms of our psyche. It might not be required.

cultural innovation, psychological adaptation becomes optional. Such a view brings into discussion one more possibility to construct evolutionary accounts of culture. It is the approach complementing the method of Tooby and Cosmides advocate, one that they are quick to reject.[37] It consists in constructing evolutionary theories about specific behavioral patterns, without making explicit their underlying psychological mechanisms. I do not advocate returning to a form of cultural adaptionsm which presupposes that all such patterns of behavior must be adaptive. Quite on the contrary. People and groups do not always succeed in addressing their problems. Also, not all of our culturally acquired behaviors must have an evolutionary utility. For instance a culturally acquired habit like smoking clearly hasn't. Therefore, adaptiveness won't constrain these outer phenomena either. Yet many social and cultural practices do effectively address problems in the environment. After all, the cultural behavior is a phenotypic trait of mankind, even one that is arguably shared by a few others species,[38] so we might expect it to be shaped, amongst other forces, by the natural selection. Also, the role of cultural behavior in the evolutionary success of mankind seems beyond doubt. Assessing the contribution each social practice or of recurring components of social practices in the survival of human groups is a worthy intellectual enterprise, one that was pursued in a broad range of disciplines, from history to human behavioral ecology, although often from a strong adaptationism stance.

This approach, mirroring that of Tooby and Cosmides, is bound to respect the autonomy of social sciences too, wile making evolutionary accounts dependent on the non-evolutionary investigations. On one hand, any evaluation of adaptedness requires prior anthropological and sociological fact-finding work to clarify the structure of the practice under scrutiny. Also, where the modeling of the relations between environmental variables and cultural practices is involved,

---

[37] Tooby and Cosmides, "The Psychological Foundations."

[38] Not only do the usual suspects, the primates, come up with techniques that they spread within the group, like for instance those very well documented in Japanese macaques (Masao Kawai, "Newly-acquired Pre-cultural Behavior of the Natural Troop of Japanese Monkeys on Koshima Islet," in *Primates* 6, 1 (1965):1-30; Ichoru Tanaka, "Matrilineal Distribution of Louse Egg-Handling Techniques During Grooming in Free-Ranging Japanese Macaques," *American Jorunal of Physical Anthropology* 98, 2 (1995): 197-201), but so do species that resemble little to man. There is strong evidence that New Caledonian crows are not only creators and employers of tools, but also that they operate changes in the design of their tools, which they socially transmit to the fellow members of their population, giving rise to divergent 'cultures' (Gavin R. Hunt and Russell D. Gray, "Diversification and Cumulative Evolution in New Caledonian Crow Tool Manufacture," *Proceedings of the Royal Society of London B* 270 (2003): 867-874).

we need the same traditional sociological or anthropological research in order to be able to test such models, by comparing the predictions with the reality on the field.[39] This makes traditional non-evolutionary research a prerequisite for any Darwinian theorizing of the human mind. On the other hand, we cannot expect culture to provide effective solutions to each problem presented by the environment. In addition, we need to leave room for cultural behaviors which lack any biological function. Both leave the selective forces unable to constrain the patterns of behavior a sociologist or psychologist can expect in a group. *Mutatis mutandis*, cultural behaviors replicate the relation of psychological mechanisms with natural selection. In sum, Darwinian inquiries into the human mind and culture turn out to be dependent of the truths established by the non-evolutionary investigations in social sciences, while the latter remain unconstrained by the former.

In other words, social sciences are autonomous with respect to evolutionary biology. But autonomous does not necessarily mean disconnected. The ties are there, but it is not the social sciences that have to take notice of what evolutionary theory has to say. On the contrary, evolutionary theory needs to employ the results of anthropology or psychology, so that it could come up with a full picture of how Darwinian forces shaped humankind, its psychological and cultural phenotype. In order to succeed, evolutionary psychology has to take into account maladaptation and contingency, which are normal parts of the life of our species, as well as the bidirectional nature of adaption. It has to fully realize that there are a myriad of historical brute facts, that cannot be anticipated by sheer adaptiveness considerations, such as cultural inventions, that shaped the way we behave and function psychologically or socially and area part of our evolutionary history. In order to fully succeed, it needs to heed to what psychology and social sciences can and need to discover independently, in order to evolutionary thinking to be able to build solid theories accounting for such realities. The sort of tight integration we expect from a properly scientific discipline with the rest of science is there, but we should somehow revise our expectations about how the ties must be forged.

[39] These models, whatever they might look, should allow for occasional maladaptive responses of groups to environmental pressures.

EPISTEMIC JUSTIFICATION, RIGHTS, AND PERMISSIBILITY

Anthony BOOTH, Rik PEELS[1]

ABSTRACT: Can we understand epistemic justification in terms of epistemic rights? In this paper, we consider two arguments for the claim that we cannot and in doing so, we provide two arguments for the claim that we can. First, if, as many think, William James is right that the epistemic aim is to believe all true propositions and not to believe any false propositions, then there are likely to be situations in which believing (or disbelieving) a proposition serves one of these goals, whereas suspending judgement serves the other, equally important goal. Second, it is in principle always possible to have different epistemic standards for evaluating the evidence for the proposition in question, so that one can have a right to believe (or disbelieve) that proposition and a right to suspend judgement on it. Whereas the first consideration counts in favour of the idea that believing justifiedly is at least *sometimes* a matter of having an epistemic right, the latter consideration favours the view that believing justifiedly is *always* a matter of having an epistemic right.

KEYWORDS: epistemic justification, epistemic rights, permissibility, Jeffrey Glick

Consider the idea that being epistemically justified in believing that *p* should be understood in terms of whether or not one has a *right* to believe that *p*. Let us call this thesis the *Right Thesis* (RT):

> RT: *S* justifiedly believes that *p* iff (i) *S* believes that *p*, and (ii) *S* has an epistemic right to believe that *p*.

As Hector Castañenda notices, RT is a view that has been advocated by several epistemologists:

> Most epistemologists, concerned with [a] person's actual beliefs, tend to understand the schematic sentence "*S* is justified in believing that *p*" as … *S* believes that *p* and *S* has a right to believe that *p* (or, *S* believes that *p* and it is permissible for *S* to do so.)[2]

[1] This paper is a collaborative effort to which each of us has contributed equally.

[2] Hector N. Castañenda, "Knowledge and Epistemic Obligation," in *Philosophical Perspectives 2*, ed. James Tomberlin (Atascadero: Ridgeview, 1988), 213. Castañenda's point seems convincing; see, for instance, Fred Dretske, "Entitlement: Epistemic Rights without Epistemic Duties?" *Philosophy and Phenomenological Research* 60, 3 (2000): 592-8, Charlotte Katzoff,

So the question of whether RT is true is independently interesting and important. But we think it is important also because of its relation to the following thesis which we call the *Permissibility Thesis* (PT):

> PT: For any person *S*, evidence base *E*, and proposition *p*, there are at least two doxastic attitudes that, given *E*, *S* is epistemically permitted to have towards *p*.[3]

As with RT, whether or not PT is true is independently important. But it is also important because of its bearing on the issue of peer-disagreement, viz. the question of whether the fact that an epistemic peer disagrees with *S* about *p* is evidence that bears on whether *p* is true and can make a difference as to whether or not *S* is epistemically justified in believing that *p*. We point this out, because there is an important relationship between PT and RT: it seems that PT entails something like RT. If two doxastic attitudes towards the same proposition *p* are equally permissible for some person *S*, then it seems correct to say that *S* has an epistemic *right* to adopt either of these attitudes. But if PT entails RT, then if there are convincing arguments against RT, PT is equally in trouble. This provides a second reason to consider whether there are convincing arguments against RT.

While there are several well-known argument's against PT, there are not many arguments explicitly directed against RT. One such argument,[4] however, goes as follows: one's having a right to φ is most plausibly understood as one's having either a *claim right* or a *privilege right* to φ. But we have neither a privilege right nor a claim right to believe certain propositions given our evidence. We, therefore, have good reason to think that RT is false. We have a claim right to φ iff all others are *prima facie* obliged not to try to stop us from φ-ing. Thus, we have a claim right to believe that *p* iff others are obliged not to try to stop us from believing that *p*. We will not discuss this conception of doxastic rights, since we are only concerned with what it is for an *individual* to be epistemically justified in

---

"Counter-evidence and the Duty to Critically Reflect," *Analysis* 60, 1 (2000): 94, and Nikolaj Nottelman, *Blameworthy Belief: A Study in Epistemic Deontologism* (Dordrecht: Springer, 2007), xii.

[3] A weaker version of PT would have it that only sometimes is it the case that there are at least two doxastic attitudes *S* could permissibly have taken toward *p* given *E*. In what follows, we focus on the stronger version of PT. We argue that there is no reason to think RT is false. This means that these arguments against RT provide no reason to think that the stronger version of PT is false. But if there is no reason to think that the stronger version of PT is false, then there is no reason to think that the weaker version of PT is false.

[4] Jeffrey Glick, "Justification and the Right to Believe," *The Philosophical Quarterly* 60, 240 (2010): 532-544.

believing a proposition, and focus on doxastic rights (rights to believe) as privilege rights.

S has a privilege right to believe that *p* iff (i) *S* is a candidate for having certain epistemic duties, and (ii) *S* has no epistemic duty or obligation not to believe that *p*. The first clause is meant to exclude those individuals, such as young children and mentally incompetent people, who seem to have neither epistemic rights nor epistemic duties. We will assume this point is correct. Next, the conception of doxastic rights as privilege rights still leaves the adherent of RT with two different options for interpreting the relevant doxastic rights. These options lead to the following definitions of RT:

RT1: *S* justifiedly believes that *p* iff (i) *S* is a candidate for having certain epistemic duties, (ii) *S* is permitted to either believe that *p* or to withhold judgement on *p*, and (iii) *S* has an epistemic obligation not to disbelieve that *p*.

RT2: *S* justifiedly believes that *p* iff (i) *S* is a candidate for having certain epistemic duties, (ii) *S* is permitted to believe that *p*, and (iii) *S* has an obligation not to withhold judgement on *p* and not to disbelieve that *p*.[5]

RT2 is clearly problematic, for on this conception of doxastic rights one has a right to believe that *p* just in case one has an epistemic obligation to believe that *p*. But if one has an epistemic obligation to believe that *p*, then suspending belief on *p* and disbelieving that *p* are impermissible, so that it seems incorrect to talk about an epistemic right in such a case.

As to RT1, Glick claims that there is a certain class of propositions such that if one's evidence justifies one in believing them, one has an epistemic obligation to believe them. RT1 says that if one is epistemically justified in believing that *p*, then it is permissible for one to suspend judgement on *p*. But then, it seems to follow that RT1 is false, for if one has an obligation to believe that *p*, it cannot be the case that one can permissibly suspend belief on *p*. Glick's crucial claim is, therefore, that we are sometimes obliged to believe some proposition. He provides two considerations in favour of this view.

First, it follows from the *Jamesian goals*. According to William James, the twin goals of cognition are to believe *all* true propositions and to believe *only* true propositions. These two goals should be in balance, for if they are not, we will either acquire large amounts of false beliefs or very few true beliefs. This means

[5] The formulations of RT1 and RT2 are our own. We think they are simpler, equivalent to, but easier to work with than Glick's JustRight$_{EPR}$I and JustRight$_{EPR}$II (see Glick, "Justification and the Right to Believe," 536-7).

that in certain circumstances one should *not* believe a particular proposition, for one would thereby run too much risk of forming a false belief. And it also means – and this is the crucial point – that in other circumstances one *should* believe (or has an epistemic obligation to believe) a particular proposition, for if one were not to believe the proposition in such circumstances one would thereby run too much risk of not forming a true belief. It follows from the twofold Jamesian goals, then, that we sometimes have an epistemic obligation to believe a proposition.

Second, it follows from cases of *overwhelming evidence*. Glick asks us to imagine that a mock jury member is presented with overwhelming evidence in favour of the guilt of the defendant. If RT_1 were true, she could conclude from the fact that she has overwhelming evidence in favour of *p*, that she is epistemically justified in believing that *p*. But then she could conclude from RT_1 that she also has an epistemic right to suspend belief on *p*. If, consequently, she withholds belief on *p*, it seems that the other mock jurors have a legitimate complaint against her: her suspension of belief seems, intuitively, epistemically blameworthy and, hence, unjustified.

However, we think that there are two ways out for the advocate of RT. First, she might point out that one could combine RT_1 and RT_2. The resulting understanding of epistemic justification is cashed out in terms of a disjunction of obligations and privilege rights and it meets the objection:

> RT_3: *S* justifiedly believes that *p* iff (i) *S* is a candidate for having certain epistemic duties, (ii) it is not epistemically permissible for *S* to disbelieve that *p*, and (iii) *S* has or does not have an obligation not to suspend judgement on *p*.[6]

We should note that there are two different ways of interpreting RT_3. On the one hand, one could think that the fact that sometimes two doxastic attitudes are equally permissible and sometimes a particular doxastic attitude is obliged means that we at least *sometimes* have an epistemic right, namely just in case two doxastic attitudes are equally permissible. On the other hand, one might think that one has an epistemic right just in case sometimes two doxastic attitudes are permissible and sometimes a particular doxastic attitude is obliged. This might seem to be an unusual definition of rights. But at least under this definition, unlike the definition alluded to in RT_2, having a right to believe that *p* does not *entail* having an obligation to believe that *p*. Nevertheless, let us work with the first interpretation for the rest of this paper.

[6] The third clause of RT_3 is, of course, logically redundant. We have put it in, however, to make the disjunct that plays such a crucial role manifest for the purposes of our discussion.

If the class of propositions on which *S* is neither obliged to believe nor obliged to suspend judgement is not empty, then epistemically justified belief is at least *in some cases* a matter of one's having a right to believe that *p*. Considering the Jamesian goals in fact provides us with good reason to think that there are such cases. This is because it seems plausible that there are cases in which one's evidence is neither clearly balanced (equally in favour of *p* as in favour of ~*p*) nor clearly in favour of *p*. In such evidential circumstances believing rather than suspending belief will favour the goal of believing all truths, whereas suspending belief rather than believing will favour the goal of not believing any falsehoods. But both goals bear equal value and one has to believe or not believe – the latter option is simply a negation of the former. This means that believing and suspending belief will be equally permissible in such scenarios. But then one has a right to believe or suspend judgement in evidential circumstances like these.

Note that RT_3 meets the objection to RT_2, for it is *not* the case that under RT_3 one is justified in believing that *p* just in case one is obliged to believe that *p*. This is because the first disjunct of RT_3's condition (iii) creates sufficient room for scenarios in which two different doxastic attitudes are equally permissible.

The second way also saves a stronger version of RT, on which to believe justifiedly *is* to have an epistemic right. It does so by claiming that what determines whether or not one's doxastic attitude is epistemically justified is not solely a matter of one's evidence bearing on *p*. It seems to us that different but equally justified epistemic standards may result in different doxastic attitudes on the basis of the same evidence base bearing on *p*, because one's evidence base for those epistemic standards may vary or one may suffer from bad luck. After all, the fact that one's epistemic standards are reliable does not imply that they are infallible. On encountering a disagreeing epistemic peer one might conclude that though her epistemic standards are reliable, she has suffered from bad luck in this case in forming a false belief (so I can consider her belief justified and yet not believe that I should alter mine). If that is true, then there will always be in principle two different doxastic attitudes that one could take that could be equally justified in the face of the same evidence base bearing on *p*, since someone else might have an equally justified set of epistemic standards, different from mine, which she uses to evaluate the evidence we both have.

One might think that for certain obviously true propositions for which we all have ample evidence, there are not and could not be any epistemic standards that would justify disbelief or withholding judgement. For instance, for the proposition *p* that there are other people we all have conclusive evidence. One might think that there could not be any epistemic standards that could justify

disbelieving or withholding judgement on *p*. But this is false. One may have good reasons to think that one frequently suffers from perceptual hallucinations, one of which is that there are other people. If so, one might have the epistemic standard that if some proposition entails that other people exist, it should not be believed. Obviously, such cases will be exceptional, but there is nothing incoherent about them. And that is all that the adherent of RT needs.

What notion of evidence should we be working with here, however? Merely one's evidence bearing on *p*? Or should we take 'evidence' to mean *all* one's evidence, one's *total* evidence, where that evidence includes both the evidence that bears on *p* and the evidence that bears on one's epistemic standards? Is it not true that *that* evidence base renders one particular doxastic attitude towards *p* uniquely justified? In other words, could not the opponent of RT just insist that when two persons *S* and *S** bring to bear different standards on their evidence in relation to *p*, then *S* and *S** do *not* share the same *total* evidence with respect to *p*, after all?

The problem we have with this move is that in making it, RT's opponent would owe us an account of what it is for one's evidence to be *total*. This turns out to be more difficult than one might initially think. First of all, as a preliminary, let us note that we are interested in what one's total evidence is with respect to the question of which evidence is relevant to whether or not one is epistemically justified in believing a proposition. Perhaps *S*'s total evidence could just be all the evidence there is, but unless *S* is some kind of deity, we do not need to bring to bear all the evidence there is in order to determine whether or not *S* is epistemically justified in believing that *p*.

The trouble is, do we allow evidence about our epistemic standards, which might include beliefs or standards about how to weigh our epistemic standards, to figure in our determination of what *S*'s total evidence is? If we do allow it, then we need to confront something like the old problem of the criterion: it seems arbitrary to stop with standards and evidence with respect to one's standards. We should include in what we consider to be our total evidence our standards and evidence employed to determine what our standards are that we employ to evaluate the evidence for one's standards, and so on, *ad infinitum*.

So how might one block the regress? It seems to us that the available strategies turn out to be incompatible with the move under consideration against RT. What all the varieties of foundationalism and coherentism do in response to the regress problem, in effect, is give up on the idea that one's epistemic standards are justified by evidence – they are 'justified' by something else. Infinitism makes a similar concession, though its proponents allow for a kind of justification –

'propositional' justification – upon which infinite evidence has a bearing. But only evidence upon which *S* has based her belief that *p* is relevant to what Infinitists call 'doxastic' justification and can thus be part of *S*'s total evidence, and only doxastic justification is, by all accounts, epistemic justification (when we are looking for what it is that justifies *S* in believing that *p*). And here is how this reflects back on PT and RT: since evidence that bears on one's standards is not evidence that is relevant to answering the question of whether *S* is justified in believing that *p*, it cannot be part of *S*'s total evidence for *p* (in relation to whether *S* is justified in believing that *p*). So it is always possible for *S* to be justified in believing that *p* and *S** to be justified in having another doxastic attitude toward *p* even where *S* and *S** share the same evidence, since *S*'s standards for evaluating the evidence might differ from those of *S**. And one cannot claim that when their standards differ, *S* and *S**'s evidence differs, without engendering an infinite regress. Or, if the opponent of RT were to make this claim, she would owe us a new solution to the problem of the criterion that did not rely on denying that epistemic standards must be justified by evidence. That is an extremely tall order, it seems to us.

We conclude that we have good reason to believe that epistemic justification should be understood in terms of epistemic rights. First, if, as William James claimed, the twofold aim of cognition is to believe all truths and not to believe any falsehoods, then there are bound to be situations in which believing (or disbelieving) and suspending belief are equally permissible, so that one has a right to believe or suspend belief in such cases. And, second, the fact that people can in principle always justifiedly hold different epistemic standards for evaluating their evidence base guarantees that there are always two epistemically permissible doxastic attitudes in any evidential circumstances, so that one has a right to believe or suspend belief. [7]

[7] For their helpful comments on earlier versions of this paper, we would like to thank Martijn Blaauw, Catarina Dutilh-Novaes, Jeffrey Glick, Fred Muller, Nikolaj Nottelmann, Herman Philipse, Jeroen de Ridder, Stefan Roski, Conor McHugh, Han van Wietmarschen, and René van Woudenberg.

UNDERSTANDING AND EXPLAINING

Jasper DOOMEN

ABSTRACT: The quest to provide a fundamental understanding and explanation of reality is an ambitious one. Perhaps it is too ambitious. The possible restrictions for such an enterprise to be successful must be inquired in order to determine the issue. Section 1 explores one's understanding in reaching (scientific) conclusions: to what extent does a successful account testify to understanding? Section 2 focuses on the other side of such an account: does it provide an explanation in a more fundamental sense than pointing out causes of phenomena, or is it restricted to such a task? A critical attitude vis-à-vis the (scientific) enterprise of unearthing reality's structure remains necessary in order not to confuse a consistent and productive theory with one that demonstrates an understanding and explanation in the sense of this article.

KEYWORDS: understanding, explaining, scientific inquiries

Introduction

In order to provide a solid basis for a (scientific) theory, understanding and explaining seem indispensable. One must understand one's findings, since otherwise the theory is nothing more than a result one has stumbled upon, as if one were to express a correct reasoning in a language one does not master, merely being able to pronounce the phones, following the syllables' sequence without knowing the meaning of the words, not being able to acknowledge the reasoning's correctness. Philosophers and scientists are supposed to have a more extensive grasp on their fields than the straw man just mentioned, not acting as mechanically as he does. Likewise, a result haphazardly reached is not said to attest to an explanation: if a result is presented, it is not sufficient that it be correct; one must also be able to make it clear *why* it is correct.

In this article, the merits of what are considered to be understanding and explaining are critically examined in that the ability to grasp a meaning may be said to constitute a necessary condition for understanding, just as the presence of an account is a necessary condition for an explanation, but that in neither situation a sufficient condition is provided for respectively an understanding and an explanation. In a modest sense, an understanding and an explanation *may* be said to occur, namely if one limits oneself to that which is empirically available. It would, however, testify to a somewhat superficial stance if one might thereby be considered to know how reality is constituted and what the fundamental reason

behind a phenomenon is (unless the possibility for such 'deeper' accounts is dismissed, a possibility that is taken seriously in this article).

In order to distinguish between the sorts of understandings and explanations, 'understanding' and 'explanation' are used in the relatively unproblematic way outlined above and illustrated in sections 1 and 2. By contrast, 'comprehensive understanding' and 'comprehensive explanation' will refer to a complete account, 'grasping' nature or reality, not limiting oneself to pointing out one or more causes of phenomena.

My position will be that comprehensive understanding and comprehensive explaining are not possible. Because I do not actually know what other (human) beings than I know, however, I must counter the objection that this perspective is too simplistic. For that reason, I will adopt a more cautious stance, and limit myself to the actions of factor-determined beings, i.e., beings whose actions are completely determined by factors. Factors are the things that determine ('make'[1]) an action if nothing else is involved. This sounds somewhat abstract, perhaps, but that is in fact unavoidable, since I cannot, being myself factor-determined, indicate which factors are actually decisive. To nonetheless illustrate the matter, presuming that an object such as a stone that is pushed down a hill is fully determined by factors such as the impulse and its shape, the factors determine the stone's path.[2]

A stone is a relatively simple object, one may say, compared with animals and human beings. Strictly speaking, I do not know whether human beings, animals or even stones are factor-determined and know only myself to be of that nature (finding no faculty within me to act alternatively from a determined way, be it on the basis of innate or empirical factors or a combination of both). Still, for the sake of convenience, I will presume the agents mentioned in the present article to be factor-determined. Actors (putatively) acting in a non-factor-determined way, on the basis of what is sometimes called 'free will,' is, for me at least, unimaginable, 'free will' only having a meaning if the freedom of movement of the will is expressed, which is, however, an idiosyncratic interpretation, 'free will' usually being used to express the agent's (as far as I am concerned incomprehensible) freedom in acting. There is (*ex hypothesi*) no way for such a being not to be factor-determined; if it should reach the same outcome as someone who comprehensively understands and/or is able to comprehensively explain, he

[1] 'Factor' originally (in Latin) means 'creator.'

[2] It is clear that this is a tautological position. That is one of the reasons why I cannot say which beings (if any) besides myself are factor-determined.

has merely stumbled upon that result and does not really find himself in the required state.

A factor-determined being would, in order to comprehensively understand, have to be able to balance the factors themselves, which is of course only possible from a factor-free position, and this is precisely what is lacking. For example, if such a being has adopted some interpretation of quantum mechanics on the basis of an education process and the independent study of authors who promote it (or oppose it, in which case this being in turn opposes *their* interpretation), the education and the result of studying are factors, as is the way the being deals with these sources of information (which may itself be the result of one or more innate or cultivated factors). This being would have to reach a temporary state of suspension of judgment with respect to these factors and then have access to the means to acquire a comprehensive explaining and a comprehensive understanding. It may – in that case, again on the basis of one or more factors – be able to doubt its knowledge and the path that has led to it, but this will be insufficient to reach the desired result; the only result that is reached is this beings' acknowledgement of its own limitations.

This little excursion, which must now, because this is not a topic to be explored here in depth, be terminated, lest the reader should be left confused with respect to the main issues to be expressed here, was merely necessary to gain some clarity on agents' position when they set out to understand and explain matters. Should they not be considered factor-determined, their acts – and therefore judgments – would be completely inexplicable (and not only in the special sense addressed in the second section of this article). That does not mean that non-factor-determined beings do not exist, of course, but only that, besides the fact that I cannot imagine their existence, for present purposes it must be assumed that factor-determined beings are the only ones that exist.

1. Understanding

Scientific theories usually receive their value upon being proved on the basis of experience; mathematical and logical theories are proved deductively. Two aspects with respect to this issue are addressed here. The first is the scientist's perspective when he validates a theory; does he 'grasp' some 'truth,' and what does this mean? This issue will be explored in the present section. The second is the proof's merit, dealt with in section 2. To differentiate, I will use 'comprehensive understanding'[3]

[3] Avoiding the more poetic but perhaps less clear alternative 'comprehensive comprehension.'

to indicate the grasp of reality and 'understanding' to indicate the ability to use theories.

What does someone's understanding something mean? Is this to be taken to reflect a 'grasp' of reality in that one knows how part of reality is constituted? I will argue that understanding rather means that one is able to utilize theories for some goal, without thereby penetrating fundamentally into reality, considering it as it is in itself,[4] whatever that may mean. Even if a theory should in some way reflect reality (in itself), it would still not be clear to the observer why reality is constituted thus and not alternatively.

The best examples to start with are logic and mathematics. What is it that one comprehensively understands (i.e., what occurs apart from being able to use the theory) if one knows that contradictions cannot occur, in line with the principle of contradiction, expressed by the formula "$\neg (p \wedge \neg p)$"?[5] It is clear[6] that one needs this information to be able to produce a valuable account at all; if one should, for example, argue that a stone that lies on top of a hill when pushed will both roll down the hill and at the same time remain where it now is, no theory that would be of use would ensue. Quantum mechanics does manifest a number of results that conflict with basic logic (e.g., Schrödinger's paradox[7]), but whether one should therefore give up some of the laws of logic or part of quantum mechanics (or at least some interpretations) – whether such a choice must be made at all depends again on the question whether one should adhere to the principle of contradiction, so the question may not have been put fairly thus – is something to be decided on the basis of other factors than comprehensive understanding (and rather by, e.g., the desire to have a consistent account).

Does someone who understands the necessary exclusion of mutually contradictory propositions comprehensively understand why this is the case? No. Their simultaneous occurrence simply doesn't work: it fails to produce viable

[4] Immanuel Kant, *Kritik der reinen Vernunft* [1781/1787]. *Kant's Gesammelte Schriften. Erste Abteilung: Werke. Band 3* (*Kant's Collected Writings. First Section: Works. Vol. 3*) (Berlin: Georg Reimer, 1904), A 42/ B 59; B 303.

[5] This principle is, incidentally, compromised (or at least not evident) on the basis of the existence of alternative views, which acknowledge the existence of contradictions (e.g., Graham Priest, "Classical Logic *aufgehoben*," in *Paraconsistent Logic. Essays on the Inconsistent*, eds. G. Priest, Richard Routley, and Jean Norman (München, Hamden, Wien: Philosophia Verlag, 1989), 141).

[6] I do not, of course, myself hereby express the occurrence of a comprehensive understanding.

[7] Erwin Schrödinger, "Die gegenwärtige Situation in der Quantenmechanik," in *Die Naturwissenschaften* 23, 48 (1935): 812.

results.[8] Besides, if one were to comprehensively understand why it is the case, alternative accounts would a priori have been refuted. In mathematics, one cannot penetrate beyond the first definitions and axioms that must be posited, such as those in Euclid's *Elements*. There is no comprehensive understanding here, either: the insights – if one wants to use that term – are not confirmed by a comparison with reality, since no such comparison is available.

This can perhaps best be illustrated on the basis of the (initial) position of Wittgenstein,[9] sometimes designated 'logical atomism,' a theory whose scope is not limited to that of mathematics, but includes it. Wittgenstein states that one must compare a picture with reality in order to come to know whether the picture is true or false.[10] Reality is the existence and nonexistence of states of affairs,[11] while the world is the whole of the *existing* states of affairs[12] (the whole of facts[13]), which entails the (idiosyncratic) position that reality comprises more than the world, a difficulty that is increased by the statement "The complete reality is the world."[14] The difference in scope between 'reality' and 'the world' is apparently not maintained here.

This contradiction is difficult to account for, but, more importantly, it is propagated that a fact and a picture must have something in common for a fact to be a picture.[15] That means that reality is approached in a somewhat procrustean manner – if something does not fit the model, it cannot be accounted for –, which is, however, not problematical as long as this model is acknowledged to be what it is: an a priori exclusion of that which cannot be expressed in language and logic.[16]

---

[8] Cf., in a different context, Richard Rorty, *Philosophy and the Mirror of Nature* (Princeton: Princeton University Press, 1979), Ch. 3, § 4, 157, 158: "The idea of 'necessary truth' is just the idea of a proposition which is believed because the 'grip' of the object upon us is ineluctable. […] The objects of mathematical truths will not *let* themselves be misjudged or misrepresented."

[9] Wittgenstein subjects his own theory to severe criticism, of course (e.g., *Philosophische Untersuchungen* [1953] Working edition, vol. 1 (Frankfurt am Main: Suhrkamp, 1997), part 1, § 114, where the contention in *Tractatus Logico-Philosophicus* that through language nature is ascertained is criticized).

[10] Ludwig Wittgenstein, *Tractatus Logico-Philosophicus* [1921], Working edition, vol. 1 (Frankfurt am Main: Suhrkamp, 1997), 2.223.

[11] Wittgenstein, *Tractatus*, 2.06.

[12] Wittgenstein, *Tractatus*, 2.04.

[13] Wittgenstein, *Tractatus*, 1.01, 2, 2.04.

[14] "Die gesamte Wirklichkeit ist die Welt." (Wittgenstein, *Tractatus*, 2.063.)

[15] Wittgenstein, *Tractatus*, 2.16; cf. 2.12: "The picture is a model of reality." ("Das Bild ist ein Modell der Wirklichkeit.")

[16] Wittgenstein, *Tractatus*, 5.4711, 5.6, 5.61.

In order to ascertain whether Wittgenstein attests to this limitation, the following is illuminating: "The state of affairs is a union of objects (things)."[17] A proposition of the form "aRb" is perceived as a picture[18] (e.g., "Colorado Springs lies to the south of Denver", if 'a' is 'Colorado Springs', 'b' is 'Denver' and 'R' is 'lying to the south of').

The problem here is that "we make pictures of facts for ourselves"[19] (the picture itself, incidentally, being a fact[20]), and in order to come to know whether the picture is true or false it must be compared with reality (cf. note 10, *supra*). So the pictures that are made of the facts are compared with reality, the positive part of which is the world (the whole of facts): the pictures made of the facts are compared with the facts. In the most antagonistic interpretation, this amounts to a circle, in which the outcome is a result of one's own contribution. In the most forthcoming interpretation, Wittgenstein's insistence that everything is experienced within the boundaries of logic and language results in a conceptual prison from which one is unable to escape. This does mean that speaking of 'the world' and 'reality' as unreservedly as Wittgenstein does is not justified (unless this is itself said to follow from one's limitations – being able to use another expression would counter these limitations); it leaves room for remarks on the 'mystical'[21] and propositions as "how the world is, is completely indifferent for what is higher. God does not manifest himself in the world."[22] The meaning of such sentences depends on how far Wittgenstein's observation is taken that the answer to the problems of life lies in their absence once the possible scientific questions have been answered.[23]

The first interpretation leads to an untenable result, 'the world,' whatever it may be in this case, remaining undisclosed. The same outcome applies in the second interpretation, but it is not equally untenable. In the latter case, the limits of knowledge are rather acknowledged, 'the world' meaning the world insofar as it can be grasped (through logic and language). In any event, logical atomism (in this guise) can merely point to one's limitations in comprehensive understanding.

---

[17] "Der Sachverhalt ist eine Verbindung von Gegenständen (Sachen, Dingen)." (Wittgenstein, *Tractatus*, 2.01.)

[18] Wittgenstein, *Tractatus*, 4.012.

[19] "Wir machen uns Bilder der Tatsachen." (Wittgenstein, *Tractatus*, 2.1.)

[20] Wittgenstein, *Tractatus*, 2.141.

[21] Wittgenstein, *Tractatus*, 6.44, 6.45, 6.522.

[22] "Wie die Welt ist, ist für das Höhere vollkommen gleichgültig. Gott offenbart sich nicht in der Welt." (Wittgenstein, *Tractatus*, 6.432.)

[23] Wittgenstein, *Tractatus*, 6.52.

To return to mathematics: it can be applied, e.g. in construction engineering. Still, that only proves the *application* and not an insight into reality: one knows what to do in order to produce a desired result, but that is all. The stability that is ever observed (the process works in this situation as it has worked in the past, because of the conformity in nature which has hitherto been present, as far as one can tell) is not itself comprehensively understood and may come to an end without the observer being able to account for such an event, something that will be dealt with in section 2. The regularity is, in other words, observed and then posited to be present in the future. This is of course no critique of the procedure that scientists follow; no engineer or doctor would be able to perform any action without resorting to such a regular pattern. I merely want to indicate that scientists do not comprehensively understand the regularity's presence and cannot therefore be assured that it will last;[24] that doesn't detract from the need to act.

In physics, the problems are even greater than in mathematics and logic. A vague notion, or rather – since not only comprehensive understanding but even understanding is absent here – word such as 'force' is used.[25] This does not invalidate the results that have been produced any more than in the cases above, but invoking notions or words that cannot be understood means that its practical outcomes constitute its sole merit. If something's cause is provided, the question 'why' it occurs or exists is not answered,[26] but merely the question 'because of what.' Indeed, Hume rightly points to a priori reasonings' insufficiency to lay bare the reason why things are as they are in matters of fact, 'cause' itself being a problematical notion,[27] a view that needs to be complemented with the position that on the basis of experience such an account cannot be produced either (which Hume, incidentally, seems to acknowledge[28]).

Physics' *applications* are not in the least struck by the present observations and these are not their focus. A lack of comprehensive understanding with its practitioners follows, however, from the mere given that they have to resort to words that only describe a process, such as 'force,' 'gravity' and 'attraction,' the

[24] Cf. David Hume, *An Enquiry concerning Human Understanding* [1748], ed. Tom Beauchamp (Oxford: Clarendon Press, 2000), 32, 33: "Let the course of things be allowed hitherto ever so regular; that alone, without some new argument or inference, proves not, that, for the future, it will continue so."

[25] This point will receive additional attention in section 2.

[26] Which is not to imply that it necessarily *can* be answered; perhaps the idea that such an answer is possible is merely a human imagination.

[27] E.g., Hume, *An Enquiry*, 60.

[28] Hume, *An Enquiry*, 26-28, 29, 30, 36.

introduction of which does not testify to a comprehensive understanding but to the fact that the limits of the ability to understand have been reached.

An even more radical account may be rendered, according to which the perspective that a theory can be said to reflect an insight (even if one grants that a comprehensive understanding is lacking) is forsaken and man's entire enterprise to make sense of the world he encounters is a mere consequence of his need to survive. In that case, he may also be said to delude himself in supposing that the problems that are solved point to a comprehensive understanding, when they are rather to be considered outcomes of an attempt to control his surroundings (an attempt that is doomed to fail, as long as there is no comprehensive understanding; any unexpected event may, after all, thwart one's plans, however carefully they may have been outlined and implemented); one may at most achieve a provisional theory, whose sole merit lies in its applicability.

2. Explaining

In a similar fashion as in section 1 with respect to 'comprehensive understanding' and 'understanding,' I will discern between 'comprehensive explaining' and 'explaining,' 'explaining' meaning that an account is provided in which one or more causes (keeping in mind the problems associated with this mentioned in section 1) for a phenomenon are discerned, and 'comprehensive explaining' meaning that the question 'why' something occurs or exists is answered. The link between the present section and the previous one is easily established. If there is no comprehensive understanding, scientific theories do not comprehensively explain anything, although they may be said to explain some phenomena, if they are successful, which is the criterion for their continuity and development.

Scientists do not proceed from a comprehensive understanding, but rather collect data on the basis of which a theory is constructed, ever in the context of the relevant background knowledge. A scientific explanation is, accordingly, fundamentally contingent, which means that it does not necessarily reflect reality; the fact that one thinks in some way does not entail that one *has* to think thus. (The word 'necessarily' is used here; of course, I cannot say that the explanation does not reflect reality sec, since this would imply a point of view on my part that is not the case, viz., that I would myself comprehensively understand and from that perspective be able to notice such a discrepancy.) For example, nature could have been constituted in such a way that objects randomly appear and disappear. It (presumably) does not behave thus,[29] at least not at the macroscopic level, but

[29] I say 'presumably' since I can only say something about nature as it appears to me.

science cannot comprehensively explain why this is the case. It can merely discern regular patterns in nature as it actually presents itself. Logic and mathematics are no less contingent in this sense, by the way, despite their claim to necessity and universality. They may apply necessarily and universally, but even if that is the case, it does not derogate from their being contingent in the present sense.

A clear sign that scientific theories do not provide comprehensive explanations is the fact that they resort to words that are merely used because the analysis cannot proceed any further. As Berkeley poignantly observes:

> That a stone falls to the earth, or the sea swells towards the moon, may to some appear sufficiently explained [by gravity]. But how are we enlightened by being told this is done by attraction? […] [N]othing is determined of the manner or action, and it may as truly (for aught we know) be termed *impulse* or *protrusion* as *attraction*.[30]

It is also important to realize that attraction is adhered to by Newton as a mathematical hypothesis rather than a "true and physical quality" ("qualitatem veram et physicam").[31]

Indeed, Newton himself insists that he does not seek to penetrate into the nature of things:

> Up to now I have exhibited the phenomena of the heavens and our sea by the force of gravity, but I have not yet pointed out the cause of gravity […] I do not contrive hypotheses. For whatever is not inferred from phenomena must be called a *hypothesis*; and hypotheses, whether they be metaphysical, physical, of occult qualities or mechanical, have no place in *experimental philosophy*. In this philosophy theorems are inferred from phenomena and rendered general through induction. […] And it is satisfactory that gravity in fact exists and acts according to the laws that have been demonstrated by us, and suffices for all motions of the celestial bodies and our sea.[32]

---

[30] George Berkeley, *A Treatise concerning the Principles of Human Knowledge* [1710] – *The Works of George Berkeley, Vol. 2*, eds. A. A. Luce and T. E. Jessop (London: Thomas Nelson and Sons, 1949), 86; cf. Hume, *An Enquiry*, 50: "There are no ideas, which occur in metaphysics, more obscure and uncertain, than those of *power*, *force*, *energy*, or *necessary connexion*, of which it is every moment necessary for us to treat in all our disquisitions."

[31] George Berkeley, *De Motu* [1721] – *The Works of George Berkeley, Vol. 4*, eds. A. A. Luce and T. E. Jessop (London: Thomas Nelson and Sons, 1951), 15; cf. Isaac Newton, *Philosophiæ Naturalis Principia Mathematica* (London: Apud Guil. & Joh. Innys, 1726), Def. VIII, 5.

[32] Hactenus phaenomena caelorum et maris nostri per vim gravitatis exposui, sed causam gravitatis nondum assignavi. […] [H]ypotheses non fingo. Quicquid enim ex phaenomenis non deducitur, *hypothesis* vocanda est; et hypotheses seu metaphysicae, seu physicae, seu qualitatum occultarum, seu mechanicae, in *philosophia experimentali* locum non habent. In

'Force' ('vis') is a vague word[33] and gravity itself is not observed,[34] the cause of a body's being brought downwards not being grasped.[35] One may, then, say that explanations are given, but no comprehensive explanations. Berkeley does not himself draw this conclusion, by the way, stating that metaphysics can provide "truly active causes" ("causae vere activae").[36]

The Newtonian theory of gravity[37] may be considered superior to that of Aristotle,[38] but only because it can account for phenomena more precisely and provide a better (in the sense of encompassing) description (or explanation) than the former. As for a comprehensive explanation, neither theory provides one, 'gravity' ('heaviness') remaining an opaque word. A comprehensively explanatory distance between one's theory and one's object may also be said to follow from the inclusion of thinking aids in one's theory that make reality a priori inaccessible, however useful such aids may be.

This can be argued for the branch of physics that deals with subatomic particles. 'Subatomic' supervenes on 'atomic,' of course ('indivisible' or, literally, 'uncuttable'). The atom is not observed but rather postulated as – in the pre-subatomic theories – the smallest possible unit. The atom is indeed postulated: "The atom is no *discovery* of natural science, but an *invention*."[39] The notion of the atom entails a contradiction, being without extension.[40] In spite of that, it is an unavoidable means.[41] The atom has fared well, for want of a better model of explanation.[42] Such models have been proposed, in the wake of the exploration of the subatomic realm. However, this development alleviates none of the potency of

hac philosophia propositiones deducuntur ex phaenomenis, et redduntur generales per inductionem. […] Et satis est quod gravitas revera existat, et agat secundum leges a nobis expositas, et ad corporum caelestium et maris nostri motus omnes sufficiat." (Newton, *Philosophiæ Naturalis*, Book 3, Scolium Generale, 530.)

[33] Berkeley, *De Motu*, 12.

[34] Berkeley, *De Motu*, 12.

[35] Berkeley, *De Motu*, 16.

[36] Berkeley, *De Motu*, 52; cf. 19.

[37] Newton, *Philosophiæ Naturalis*, Def. V, 3, 4; Book 3, Regula 3, 388, 389.

[38] Aristotle, *Physica* [± 350 BCE] – *Aristotelis Opera, Vol. 1*, ed. I. Bekker (Darmstadt: Wissenschaftliche Buchgesellschaft, 1960), Book 8, 255b.

[39] "Das Atom ist keine naturwissenschaftliche *Entdeckung*, sondern eine *Erfindung*." (Hans Vaihinger, *Die Philosophie des Als Ob* (Leipzig: Felix Meiner, 1922), 150.)

[40] Vaihinger, *Die Philosophie*, 102, 605.

[41] Vaihinger, *Die Philosophie*, 104, 105.

[42] Vaihinger, *Die Philosophie*, 450.

Vaihinger's remark that "*without* the atom, science collapses, but, to be sure, true knowing and understanding is impossible *with* it."[43]

The value of the carefully constructed atomic and subatomic models must be acknowledged, but also the corollary of their introduction, namely an absence of comprehensive explanation.[44] This entails that a 'theory of everything' in the sense of a theory that aims at unlocking "[…] the cardinal secrets of nature so as to render physical reality comprehensively intelligible,"[45] may consist in an account that explains all phenomena (or at least all physical phenomena), but not in a definitive theory in the sense that it would offer a comprehensive explanation for all aspects of reality.[46] After all, physics is characterized by an empirical approach no less than the other sciences, collecting data and constructing a theory by integrating them into a meaningful synthesis.[47] For a comprehensive explanation,

---

[43] "*Ohne* das Atom fällt die Wissenschaft; aber allerdings – wahres Wissen und Erkennen ist *mit* demselben nicht möglich." (Vaihinger, *Die Philosophie*, 102.)

[44] Fine's conclusion is more radical than mine: "If pressed to answer the question of what, then, does it *mean* to say that something is true (or to what does the truth of so-and-so commit one), NOA [the natural ontological attitude] will reply by pointing out the logical relations engendered by the specific claim and by focusing, then, on the concrete historical circumstances that ground that particular judgment of truth. For, after all, there *is* nothing more to say." (Arthur Fine, *The Shaky Game. Einstein, Realism and the Quantum Theory* (Chicago, London: University of Chicago Press, 1986), 134.)

[45] Nicholas Rescher, *Nature and Understanding. The Metaphysics and Method of Science* (Oxford: Clarendon Press, 2000), 74.

[46] As Duhem observes, "Explaining, *explicare*, is to reveal the *reality* of *appearances* that cover it like veils, in order to see this reality stripped and face to face. The observation of physical phenomena does not put us in touch with reality hidden behind the sensible appearances, but with these sensible appearances themselves, taken in a particular and concrete form. The experiential laws no more have material reality for their object; they deal with these same sensible appearances, albeit taken in an abstract and general form." ("Expliquer, *explicare*, c'est dépouiller la *réalité* des *apparences* qui l'enveloppent comme des voiles, afin de voir cette réalité nue et face à face. L'observation des phénomènes physiques ne nous met pas en rapport avec la réalité qui se cache sous les apparences sensibles, mais avec ces apparences sensibles elles-mêmes, prises sous forme particulière et concrète. Les lois expérimentales n'ont pas davantage pour objet la réalité matérielle; elles traitent de ces mêmes apparences sensibles, prises, il est vrai, sous forme abstraite et générale.") (Pierre Duhem, *La Théorie Physique. Son Objet et sa Structure* (Paris: Chevalier & Rivière, 1906), 6.) Physical theories do not provide a comprehensive explanation (Duhem speaks of 'explanation' (so without 'comprehensive'), of course) (Duhem, *La Théorie Physique*, 26, 38; cf. 171, 361, 362).

[47] One may even be more critical and say that such a methodology cannot even lead to an explanatory account, a holistic theory being necessary for such a result (Rescher, *Nature and Understanding*, 78-80).

another model would be needed as a necessary condition; whether it would also be a *sufficient* condition would depend on the sort of model.

Those who aspire to establish a 'theory of everything' as an alleged reflection of the structure of reality seem, then, to be confined to the metaphysical stage as Comte describes it:

> In the metaphysical state […], the supernatural agents are replaced by abstract forces, veritable entities (personified abstractions) inherent in various beings of the world, and conceived as capable to engender by themselves all observed phenomena, the explanation of which consists, then, in assigning to each the corresponding entity.[48]

Indeed,

> the final term of the metaphysical system consists in conceiving, instead of multiple particular entities, a single great entity, *nature*, considered as the unique source of all phenomena.[49]

Once the level of application is considered, the same analysis pertains. From the fact that one knows how to reach a desired result, no comprehensive explanation follows.[50] If a doctor manages to treat a patient successfully, or even cures a disease, all he does (which is not to imply that this is a slight task) is to combine several data to find one or more causes of a disease that can subsequently be abated or removed. An appeal to notions that cannot be further elucidated, such as 'cell' or 'gene,'[51] remains necessary. Moreover, from the observation that certain behavior is prone to lead to a disease, or, conversely, its absence, no

---

[48] "Dans l'état métaphysique […], les agents surnaturels sont remplacées par des forces abstraites, véritables entités (abstractions personnifiées) inhérentes aux divers êtres du monde, et conçues comme capables d'engendrer par elles-mêmes tous les phénomènes observés, dont l'explication consiste alors à assigner pour chacun l'entité correspondante." (Auguste Comte, *Cours de Philosophie Positive, part 1* [1830] – *Œuvres d'Auguste Comte, Vol. 1* (Paris: Éditions Anthropos, 1968), 3, 4.

[49] "[…] le dernier terme du système métaphysique consiste à concevoir, au lieu des différentes entités particulières, une seule entité générale, la *nature*, envisagée comme la source unique de tous les phénomènes." (Comte, *Cours,* 4.)

[50] Cf. Richard Rorty, *Objectivity, Relativism, and Truth. Philosophical Papers Vol. 1* (Cambridge: Cambridge University Press, 1991), 5: "On an antirepresentationalist view, it is one thing to say that a prehensile thumb, or an ability to use the word 'atom' as physicists do, is useful for coping with the environment. It is another thing to attempt to *explain* this utility by reference to representationalist notions, such as the notion that the reality referred to by 'quarks' was 'determinate' before the word 'quark' came along […]."

[51] Cells' and genes' internal structures can of course be uncovered, but that does not lead to a comprehensive explanation.

answer to the question 'why' that is the case ensues, and one remains in the realm of explanations in the sense of 'because of what' as mentioned in section 1. All that is provided in medicine is a generalized observation. Even if some cure exhibits universal results (the cure being effective in each instance encountered hitherto), one still has not comprehensively explained why. A comprehensive explanation would bring with it that doctors could not be surprised by a new case in which the cure would prove not to be effective (in which case a lack of universality would in hindsight be established), but that cannot be guaranteed in medicine,[52] some new situation that had not been considered being ever possible.

Doctors only discern a regular pattern, oblivious why certain diseases occur in certain cases, only able to observe causes. Apparently, some behavior leads to a disease; somewhere, the explanation ends (so that no comprehensive explanation is given), the difference with previous theories being that one is now able to give a better explanation in the sense that the deepest cause one can find is further removed from the surface than the deepest ones that appeared before, evidenced in doctors being better equipped to combat illnesses than their precursors were. For medicine's purposes, a comprehensively explicative account may not be required (although its presence would probably be welcomed), but it does question the justification of placing this discipline on a pedestal.[53] Of course, Molière's discrediting of doctors, *inter alia* on account of their (obviously circular) appeal to a 'dormitive virtue' ('virtus dormitiva') to explain why opium makes someone sleep,[54] is not fully pertinent, at least not anymore, especially since they do know how to cure some patients, as opposed to those derided by him for not being able to do so.[55]

Presuming that animals (all varieties, from ants to chimpanzees) are factor-determined beings and can be said to use their abilities for survival purposes only,

[52] Or other fields of research, for that matter.

[53] Cf. Richard Rorty, *Contingency, Irony, and Solidarity* (Cambridge: Cambridge University Press, 1989), 8: "From our point of view, explaining the success of science, or the desirability of political liberalism, by talk of 'fitting the world' or 'expressing human nature' is like explaining why opium makes you sleepy by talking about its dormitive power. To say that Freud's vocabulary gets at the truth about human nature, or Newton's at the truth about the heavens, is not an explanation of anything. It is just an empty compliment – one traditionally paid to writers whose novel jargon we have found useful."

[54] Molière, *Le Malade Imaginaire* [1673] – *Œuvres de Molière, Vol. 9*, eds. E. Despois and P. Mesnard (Paris: Librairie Hachette, 1886), 443.

[55] Molière, *Le Malade*, 397.

acquiring no comprehensive explanations (and not even explanations, perhaps[56]), this consideration may be extended to human beings (if they are factor-determined beings), so that any theory they constitute does not represent reality, but only serves to (very modestly) control nature.[57] This approach may, in the light of what has been said hitherto, be more convincing than one which does adhere to such a representation.[58] The theories that are proposed and corroborated may be ever so intricate and impressive, that takes away nothing of the divide that separates them from a comprehensive explanation, which would only be possible if another approach than the prevailing one were available. The difference between a description and an explanation may, accordingly, be said to be gradual, if such a difference can be upheld at all.

3. Consequences

It appears that no attempt to gain a comprehensive understanding or a comprehensive explanation has been successful. The two appear to be interrelated in that the acquisition of one entails the other. Whether this is indeed the case remains, strictly speaking, a matter of speculation until the stage of comprehensive understanding or comprehensive explanation is reached. One may, however, doubt the possibility of such a result, not only whether this is feasible for factor-determined beings but whether it is possible at all. Are endeavors to gain a comprehensive understanding as a more fundamental understanding and a comprehensive explanation as a more fundamental explanation than those provided by the sciences not a priori doomed to fail, since they are directed at something that is not available, from any perspective whatsoever?

Perhaps that is the most viable way to approach these issues. Perhaps, then, it must be said, with James: "Purely objective truth, truth in whose establishment the function of giving human satisfaction in marrying previous parts of experience with newer parts played no role whatever, is nowhere to be found. The reasons why we call things true is the reason why they *are* true, for 'to be true' *means* only

[56] I say 'perhaps,' since this depends on the scope of the notion 'explanation'; it may be man's prerogative to explain matters, animals being unable to do so, but only if reason is supposed to be a special faculty vis-à-vis the other faculties inherent in man and the animals rather than the apex (as far as I can tell) of a hierarchy of skills to (modestly) control nature.

[57] Cf. Rescher, *Nature and Understanding*, 134-140.

[58] In addition, the idea that phenomena can be described and explained in a single, correct way, one creature having the privilege of being capable to do so, may be deemed "merely mythology" (Rescher, *Nature and Understanding*, 131).

to perform this marriage-function."[59] Whether this is correct cannot be said by a factor-determined being, since it would otherwise already have transgressed its conceptual limits and have entered the realm that is, *ex hypothesi*, undisclosed to it, and whether other beings than factor-determined ones (can) exist is just as speculative, at least for a factor-determined being, let alone the answer to the question whether those beings would be able to reach such a state if they did in fact exist.

For the practical sciences, this result has but few consequences, as was remarked specifically with regard to comprehensive explaining in section 2. For those sciences that aim to unravel nature's secrets, however, a need to reflect on the validity and possibility of their pursuit arises. Unless a means to construct an alternative method to the one prevalent in the scientific method, viz., a construction of a theory on the basis of empirical input, a model that works quite well in practice but provides neither a comprehensive understanding nor a comprehensive explanation (the contents of which model I am of course unable to provide, not even being able to indicate whether such a model is possible at all), is found, the realm of comprehensive understanding and comprehensive explaining is unattainable. Such a result is obviously unsatisfactory, but the only one that can warrantably be said to follow from the foregoing analysis.

It is tempting to say that a domain of comprehensive understanding and a domain of comprehensive explaining are unreachable a priori since they are illusions, fantasies created to have something to aspire to, and that the regular patterns scientists discern are all there is. First of all, this means the acknowledgement of science's limitations. Second, as I mentioned, this is, from the viewpoint of a factor-determined being at least, just as speculative as positing such a realm. One is unable to determine whether nature holds great secrets (whether they be ultimately inaccessible or not) or rather merely presents the material to fabricate the illusion that such secrets would exist, just as it is impossible (for now at least) for a factor-determined being to grasp such secrets if they do in fact exist. In any event, it is unwarranted to identify discovering a regular pattern in data with comprehensive understanding and comprehensive explaining; all this points to is a regularity, the basis of which remains elusive if all one is able to do is observe it. As long as scientists' activities are limited to induction (or, in mathematics in logic, deduction), however intricate their pursuits may be, no comprehensive understanding or comprehensive explaining is realized.

---

[59] William James, *Pragmatism. A New Name for Some Old Ways of Thinking* [1907] (Cambridge, London: Harvard University Press, 1975), 37.

Conclusion

The foregoing easily leads to the conclusion that one is delivered to a forlorn skepticism. No fundamental, comprehensive understanding or comprehensive explanation of that which is encountered is provided. Since this situation is inescapable for a factor-determined being, won't an unbridgeable chasm to reality (or nature) remain forever? That depends on one's position. The very notion of a realm of reality existing independently of reasonable inquirers, waiting to be discovered, understood, explained or – according to others than myself – comprehensively understood and comprehensively explained, may have to be relinquished.

I cannot myself draw this conclusion, nor aver the opposite, as I am, after all, a factor-determined being. Strictly speaking, then, some degree of skepticism remains. This is not problematical in practice for most sciences, whose practitioners will unencumbered continue their pursuits, and whose successes are undisputed, as long as they provide actual results. Those sciences that aim at comprehensively explaining reality, however, will need to reflect the very possibility of such a goal. It cannot a priori be said to be fruitless – also because I am factor-determined –, but considering the (necessarily empirically uncorroborated) notions they have smuggled in (or, less unfavorably, posited), the need to temper their ambitions appears to be a given.

JUSTIFICATION AND TRUTH CONDITIONS IN THE CONCEPT OF KNOWLEDGE

Dale JACQUETTE

ABSTRACT: The traditional concept of propositional knowledge as justified true belief (JTB), even when modified, typically in its justification condition, to avoid Gettier-type counterexamples, remains subject to a variety of criticisms. The redefinition proposed here puts pressure more specifically on the concept of truth as redundant in light of and inaccessible beyond the most robust requirements of best justification. Best-J is defined as justification for believing in a proposition's truth where there is no better countermanding justification for believing instead the proposition's negation. A pragmatic perspective argues that truth is unnecessary and unattainable as a condition of knowledge beyond the requirement for practically attainable best justified belief. The key argument with respect to the eliminability of the truth condition in favor of a properly tailored justification condition is that there is nothing we do or can do in trying to satisfy the truth condition for knowledge beyond considering the epistemic merits of the justification that a believer accepts in coming to believe that the proposition is true.

KEYWORDS: Gettier problem, epistemic justification, knowledge, pragmatic epistemology, truth

Ob ich etwas *weiß*, hängt davon ab, ob die Evidenz mir recht gibt, oder mir widerspricht. (Whether I *know* something depends on whether the evidence backs me up or contradicts me.) — Wittgenstein, *Über Gewißheit* §504

1. Knowledge as Justified True Belief

The conceptual analysis of propositional knowledge as justified true belief (JTB) originates with Plato's dialogues *Meno* and *Theaetetus*.[1] The definition admittedly captures something essential to the concept of knowledge, but remains problematic in any recognizable Platonic formulation, even when it is modified, generally by qualifying justification condition J as a reformulated J*, in a fortified (J*TB) analysis intended to avoid Gettier-type counterexamples.[2]

[1] Plato, *Meno* 97e-98a; *Theaetetus* 201d-e.

[2] Edmund L. Gettier, "Is Justified True Belief Knowledge?" *Analysis* 23 (1963): 121-123. Roderick M. Chisholm typifies the reaction to Gettier's counterexamples to the traditional analysis of the concept of knowledge in strengthening the justification condition in JTB to 'nondefective'

We may motivate the introduction of a justification condition in analyzing the concept of knowledge, not by invoking Socrates' metaphor of true belief as a valuable statue that runs away if it is not chained down by good reasons, but by considering epistemic scenarios in which a proposition is believed that may also happen to be true, but that intuitively does not constitute knowledge. A gullible person who sincerely believes what he or she is told by a fortune teller, even if the proposition turns out to be true, is not generally regarded as having knowledge. Something more is needed, a distinctively epistemic requirement that we should support our belief with justification, proof, evidence, warrant, reasoning, and the like, in order to know that a given proposition is true.

A major difficulty in the Platonic JTB or J*TB definition of knowledge is the inclusion of a truth T condition independently of the best justification of which we are capable, and of the comparatively weak, even when Gettier-proofed, unqualified justification condition J or J*. I propose eliminating truth T as a *condition* of knowledge altogether in favor of a strengthened *best justification* (Best-J) condition, defined as a humanly attainable requirement to provide a standard for a belief's constituting knowledge in a strengthened Best-JB definition of knowledge. By the proposed definition, the best justified (Best-J) beliefs are beliefs in the truth of whatever propositions we are (a) justified in believing to be true, (b) when there is no better countermanding justification for their negations. We can think of Best-J as a suggested replacement for both J or J* and T in an analysis in an analysis of propositional knowledge. Among its other virtues, as explained below, Best-J offers an intuitively satisfying forestallment of Gettier counterexamples. The idea is not merely to replace J*TB by Best-JTB, but rather by Best-JB, eliminating T altogether, and thereby effectively reducing the contribution of the truth condition in the traditional and Gettier-proofed definitions of knowledge to best justification condition Best-J. The account preserves truth as a *concept*, even if it does not make truth a *condition* of knowledge. It is moreover compatible with the anti-skeptical expectation that truth is a potentially attainable goal for epistemic justification. If we choose, we can preserve the traditional concept of truth in an analysis of knowledge as Best-JB. We are free to suppose that truth is nothing more than the descriptive aptness of a proposition linguistically representing a corresponding existent truth-making state of affairs. Truth is nevertheless a semantic concept, rather than epistemic in

justification. Chisholm, *Theory of Knowledge*, 3rd edition (Englewood Cliffs: Prentice-Hall, 1989, previous editions 1966, 1977), 90-99.

the usual sense; so arguably, as others have also charged, truth philosophically has no business as a condition of knowledge.[3]

Significantly, in our actual epistemic practice, we do not invoke the transcendental truth of our beliefs as a condition of knowledge. We do not need to treat our knowledge of truth as existing independently of, and as though we had direct access to what is true and what is false above and beyond our judgments as to which propositions admit of the best available justification. Justification does the heavy lifting in discovering knowledge and supporting knowledge claims. Truth as a property of propositions transcending what we can learn from the best justification practically available to us is a condition that can only be satisfied by a godlike transcendent intelligence. If we want to bring epistemic theory into line with epistemic practice, then arguably we should give up justification-independent and justification-transcending truth as a requirement for knowledge in favor of another condition that accomplishes the same purpose as that of best justification. By trading in JTB or J*TB for Best-JB, we bring epistemology back down to earth, eliminating truth as a condition of knowledge independently of the best justification of which we are capable, but without stepping away from the concept of truth as an attainable goal in the pursuit of best justification. Then we can actually acquire knowledge of as well as belief in true propositions, for a recommended use of 'knowledge' and its cognates in a nonredundant subcategorization of true as well as best justified belief. Part of the point is that we also thereby avoid *epistemic hypocrisy*, failing conscientiously to practice what we preach in demanding truth of knowledge while relying entirely on assessments of best justification to settle all questions as to a proposition's truth.

2. Knowledge and Knowledge Claims

John Hick writes in *Faith and Knowledge: A Modern Introduction to the Problem of Religious Knowledge*: "knowledge cannot (by definition) be erroneous; but it is always possible for a knowledge claim to be erroneous."[4]

---

[3] Ansgar Beckermann makes this point eloquently in "Wissen und wahre Meinung," in *Die weite Spektrum der analytischen Philosophie: Festschrift für Franz von Kutschera* (Berlin: Walter de Gruyter, 1997), 24-43. Jay F. Rosenberg quotes and translates the most telling passages from Beckermann, p. 42 in his *Thinking About Knowing* (Oxford: Clarendon Press, 2002), 202-207. That Beckermann's position remains at odds with both Peircean pragmatic epistemology and the analysis of knowledge as Best-JB is indicated in this selection, p. 41 (cited in Rosenberg, *Thinking About Knowing*, 204): "Truth is the goal, and justification only a means or a criterion. What we aspire to are *true beliefs*. Whether a belief is *justified* interests us only because as a rule its truth is not obvious."

We distinguish between being epistemically justified in making a *knowledge claim*, in circumstances where we have the right epistemically to claim that we know something, and the *content of the knowledge claim* itself being epistemically justified. It appears that we can only make progress toward deciding what is actually known and what is only believed or claimed to be known on grounds of best justification. Pragmatically speaking, the way we actually validate knowledge claims in practice is by appealing to whatever we take to be the best justification for the truth of whatever propositions are supposed to be known. If the evidence bears out the knowledge claim, as Wittgenstein maintains in the motto quoted at the beginning of the essay, then we judge it to constitute genuine knowledge. If not, then we withhold classifying the belief as something that is actually known.

As we define best justification, it is relative ultimately to the available explanations and the state of scientific findings and method at a particular time. There are cultural historical conditions both for being justified in making a knowledge claim and in judging whether or not a knowledge claim is true, whether or not the conditions for genuine knowledge are actually satisfied. Thus, I may be epistemically justified in 1830 in claiming to know that space is Euclidean, rectilinear, and infinitely extensive and divisible, even if the proposition considered more timelessly or in light of today's improved relativity physics based on non-Euclidean geometry, is no longer epistemically justified. Responsible epistemic practice dictates that finite epistemic agents like ourselves should avoid claiming access to a proposition's transcendental truth, independently of whatever best justification we are actually capable of achieving, in order to support the belief that a proposition for which knowledge is claimed is actually true.

We seek knowledge by trying to justify a proposition's truth or falsehood (the truth of its negation), in the course of which we strive to accept only the most strongly justified beliefs. When we are confident in our ability to provide convincing epistemic justifications for our findings, then we publish knowledge claims to the effect. If there is no stronger countermanding evidence for the negation of what we claim to know, then we are justified at least in making such knowledge claims and judging them to be true. We never break outside the bounds of knowledge *claims* anyway, even if with excellent justification we claim to know that we know that a relevant proposition and corresponding knowledge claim is true. If stronger countermanding evidence should arise, as judged, for example, by the evolving standards, theories, methods and instrumentalities of

[4] John Hick, *Faith and Knowledge: A Modern Introduction to the Problem of Religious Knowledge*, second edition (Eugene: Wipf & Stock Publishers, 2009), 208.

epistemically self-improving science, then it may turn out that an original knowledge claim is false after all, that the subject did not possess genuine knowledge, and that our judgments upholding the truth of the knowledge claim at the time have also failed to be supported in the long run by what later appears to be better justification.

If we are serious about the practical requirements of knowledge and knowledge claims, and hence about the available justifications for each, then we should not lose sight of the epistemic fallibility of finite epistemic agents. What any finite epistemic agent believes to be true, no matter how ardently, authoritatively or forcefully expressed, need not actually be the case. Absolute transcendent truth, independent of the best justification of which we are capable, is a condition suitable only for a godlike subject's knowledge. We approach truth by managing our best justified beliefs expressed also as knowledge claims. If immediate access to justification-transcending truth were within the reach of any cognitive subject, the effect would be to make any justification condition obsolete in defining the concept of knowledge, reducing knowledge to true belief. An ideal godlike epistemic agent with direct justification-independent access to a proposition's truth never stands in need of any type of justification in making and assessing the truth or falsehood of knowledge claims.

3. Fallibility and Pragmatically Best Justification

An alternative to JTB and J*TB analyses of the concept of knowledge involves a more exact definition of best justification or Best-J. The concept is intended to provide a strengthening or qualification of justification condition J in the traditional Platonic definition of knowledge, and of condition J* in beefed-up J*TB variations designed to forestall Gettier counterexamples to JTB. Like the original analysis, these sometimes kludgy stop-gate efforts still require truth T as a condition for knowledge.

To avoid epistemic hypocrisy, to bring epistemology back down to earth, to make knowledge claims more susceptible of confirmation and disconfirmation, and for the sake of still further advantages, we eliminate truth as a condition of knowledge and replace it with best justification, defined in this way:

> Best-J: Doxastic subject *S* is best (albeit defeasibly) justified (Best-J) at time *t* in believing proposition *p* =df (a) *S* is justified at time *t* in believing proposition *p*, and (b) there is at *t* no countermanding better or stronger justification available in practice for any doxastic subject to disbelieve proposition *p* or any proposition invoked in justifying belief in the truth of proposition *p*, or to believe instead the negation of proposition *p* or at least one proposition invoked in justifying belief in the truth of proposition *p*.

Best justification is the best epistemically that we can do, and therefore the best that we should be expected to do. The revised analysis of the concept of knowledge as Best-JB depends on the best justification for belief in a proposition's truth. It demands strong justification for a subject to believe that the proposition is true, if the belief is to constitute knowledge and for the corresponding knowledge claim to be correctly judged true, and where there is no better justification practically available at the time to support the contrary evaluation that the relevant proposition is false.

Accordingly, we do not need to speak of best justification in a sense requiring successively more tests or iteratively collaborative evidence of a proposition's truth in an endless pursuit of the absolutely 'best' justification. Such expectations can only be associated instead with another more ideal sense of the concept than we propose. We consider instead the qualitative condition of satisfying the highest prevalent standards of justification practically available to a subject in arriving at a knowledge claim. We are then invited to entertain our own knowledge claims or meta-claims about the truth or falsehood of these knowledge claims. The point is that Best-J must actually be practically available to the epistemic subject in issuing or validating a knowledge claim. Impractical extremes of justification imagined are excluded by the concept of best justification. We say only that a proposition is best justified if it is justified in the usual sense and there is no better justification for its negation. That is a demanding but still attainable condition that we can and often do satisfy in our practical knowledge verifying and amplifying activities.

We decide upon the best epistemic justification in practice as whatever we deem to provide the maximally practically attainable reason for accepting a proposition as true. Otherwise, we risk the possibility that there may exist better justification for the proposition's negation. We are well advised in seeking Best-J beliefs to take as our guide the most strongly corroborated work in observational and experimental science and proto-science. The concept of truth is needed for the concept of knowledge, even if truth is eliminated as a condition of knowledge, since truth is plausibly identified as the goal at which epistemic justification aims.

The revised Best-JB (minus T) analysis of the concept of knowledge must itself be true if it is to have any ultimate philosophical significance. But as for any other knowledge claim, we are not required to establish its truth independently of its corroboration by the best justification of which we are capable, in order to know, if and when we reach that point, that the Best-JB analysis of the concept of knowledge is adequate. A pragmatic perspective argues that transcendent truth is unnecessary and unattainable as a condition for knowledge anyway, because, as

the terminology indicates, it looks beyond the requirements for a best or maximally practically attainable justified belief. What from a pragmatic perspective we do not strictly need in knowledge theory, on the contrary, is the pretense of applying a justification-independent, justification-transcending condition of absolute truth in order for a belief to constitute genuine knowledge.

4. Argument for Truth in Knowledge

It might nevertheless be objected that truth is strictly needed for the concept of knowledge. The argument is that if we do away with the truth condition, then by default we allow false propositions to count as knowledge.

The criticism can be answered in several ways. First, we emphasize that, when in doubt about a proposition's truth, rational epistemic agents inevitably appeal to the best justification available for a proposition's truth or that of its negation, and not to any justification-independent grasp of the proposition's justification-transcending truth. We might say as shorthand in rejecting a knowledge claim that the proposition in question is not true. What we mean by this, upon consideration, and if we are entitled to assert it at all, is generally that there is no best justification for the proposition. In that case, then, there is either nothing we are willing to count as justification for the proposition, or there is better countermanding justification instead for the proposition's negation, in support of its negation.

The fortune teller example is a good case in point. We visit the fortune teller, who, thanks to our gullibility, instills in us a belief in the truth of a certain proposition that turns out actually to be true, but which does not seem to constitute knowledge. Under traditional JTB, the fortune teller case ought to count as knowledge, because there is a true belief for which there is justification in the watery sense that the fortune teller's presumed authority offers a reason to believe. It is just not very good justification or a very good reason for accepting the belief. The traditional JTB definition seems to imply that fortune teller inspired belief is knowledge, because it does not discriminate between different kinds or degrees of justification, but requires the knower only to be in possession of 'justification' without further qualification. The alternative is to argue that consulting a fortune teller does not constitute justification of any kind or in any sense for belief in a prophesied proposition's truth. To exclude fortune telling as epistemically justificatory is to rely on potentially controversial substantive background assumptions contradicting the thesis that fortune tellers have knowledge of the future. More significantly, it is also to depart from the informal description of epistemic justification as having a reason for believing a

proposition's truth. For the gullible person in the relevant sense has a reason for believing what the fortune teller says. If you ask the gullible believer, "Why do you believe that?" or "What is your reason for believing that?," the gullible believer will answer, "Because the fortune teller told me it would happen." The answer is likely to be widely accepted at face value as an epistemic justification, and not misconstrued as an effort at identifying the belief's origin or cause. It is a reason for the gullible person to believe, although it is evidently not a very good reason; a justification, but not a good justification and certainly not the best.

Is there then better justification for the negation of whatever proposition the fortune teller has induced the gullible believer to accept? The way in which common sense treats the example is to argue that fortune tellers generally are unreliable in forecasting the future. When they get it right, in any case, there is no connection between whatever hocus-pocus they perform and the state of the world from which knowledge of a future state might be more reliably predicted. If some people could simply 'see' into the future, finding patterns in tea leaves or in the depths of a crystal ball, then our negative assessment of the justification status of fortune telling might be softened. Since we do not believe this on independent grounds, we generally reject fortune telling as epistemically justificatory. If we consider the chain of reasons by which the gullible believer tries to justify the fortune teller's deliverances, then, even if the proposition which the fortune teller foretells happens to be true, we reject the mere fact that a fortune teller has made such a pronouncement as Best-J best justification for the proposition's truth. The gullible believer's chain of reasoning is this:

(1) The fortune teller said that a certain future event E will occur.

(2) Fortune tellers reliably accurately predict the future.

(3) I am justified in believing that event E will occur, more or less as the fortune teller said.

The epistemically weak but inferentially indispensible link in the chain is (2). If we think that there is better justification for the negation of (2), for the proposition instead that fortune tellers do not reliably accurately predict the future, then we will have cut the ground from under the gullible believer's justification. In criticizing the fortune teller example we are not obligated to raise doubts about the best justification of the whatever it is that the fortune teller has prophesized, which we agree all along will turn out accidentally to be true. We will then have shown that there is better justification for the negation of something essential to what the gullible believer believes by reference to which belief in the content of the fortune teller's pronouncement is finally supposed to

be justified, and hence by extension that the gullible believer lacks Best-J best justification for believing in the truth of the hypothetically but accidentally true proposition that event E will occur.

Appealing to best justification is always good enough in making and judging knowledge claims. We can do no better in practice when questions of truth arise. We seek and can reasonably expect no other arbiter of whether or not a given proposition is true than whether or not it satisfies a best justification condition. This makes the T truth condition in JTB and J*TB not only redundant but pretentious and hypocritical in light of the Best-J condition. Unlike ideal godlike epistemic agents, we finite thinkers, independently of the best methods of justification, have no direct access to the justification-transcending truth of propositions involved in justifying and evaluating knowledge claims. The semantic truth condition T in JTB, or the Gettier-resistant J*TB analysis of the concept of knowledge, is objectionable as well because it does no distinct work apart from that shouldered by a properly interpreted and properly applied epistemic justification condition. We address these difficulties in the proposed Best-JB analysis of the concept of knowledge by substituting the best justification Best-J condition for both unqualified J or Gettier-resistant justification condition J* and truth condition T in JTB and J*TB. We defend the pragmatic Best-JB analysis as offering significant improvements over ideal Platonic-Socratic JTB and Gettier-proofed J*TB definitions of the concept of knowledge.

5. Advantages of Best-JB over JTB and J*TB

Altogether, we can call upon at least eight theoretical advantages of Best-JB over JTB and J*TB in defending the analysis of propositional knowledge as Best-JB. We consider the following reasons as contributing to the philosophical case for Best-J and Best-JB, by virtue of: (1) Ockham's razor. (2) Avoiding epistemic hypocrisy in theory and practice. (3) Projecting a practically attainable ideal of best justification, and hence of knowledge according to the analysis, bringing epistemology pragmatically back down to earth. (4) Making justification scientific. (5) Avoiding what we shall call flimsy Borgesian 'anthill' justifications. (6) Avoiding Gettier counterexamples without *ad hoc* provision. (7) Explaining reversals of knowledge claim validations. (8) Offering at least an equally good solution to the problem of universal ignorance when compared with condition T in JTB and J*TB.

Advantage 1: Ockham's Razor

By eliminating truth condition T from JTB and J*TB, and replacing justification condition J or J* with best justification condition Best-J, the Best-JB analysis of the concept of knowledge is conceptually more economical and arguably simpler in its analysis and application. We aim at truth in knowledge-seeking, even when it proves to be a moving target. For finite epistemic agents the implication is that, informally speaking, we can only seek the best justification for a proposition's truth of which we are practically capable in our historical circumstances. Knowledge, in the sense of what we are best justified in claiming to know, is made relative in this way to the developmental state of our science and philosophy of scientific methodology. We can be justified in making a knowledge claim under these circumstances when we believe ourselves to be in possession of best justification for a particular proposition, even if it should eventually turn out that our belief that we know and what it is that we believe or claim ourselves to know are not sufficiently supported by what turns out afterward to be best justification, when scientific methods of justification are improved over time.

Advantage 2: Avoiding Epistemic Hypocrisy

By replacing conditions J and T in JTB (and J* in J*TB) with Best-J in Best-JB, we also avoid epistemic hypocrisy. Epistemic hypocrisy is preaching something other than we practice in defining or otherwise explaining what it takes for a belief to constitute knowledge. If we preach *ex cathedra* that *truth* is a *condition of knowledge*, but our actual epistemic practice does not involve direct access to justification-independent, justification-transcending truth, then we are guilty, as I define the phrase, of epistemic hypocrisy.

Best-JB avoids epistemic hypocrisy by eliminating justification-transcending truth condition T from the analysis of the concept of knowledge and replacing it along with J or J*. It does so in principled recognition of the fact that knowledge for pragmatically-contexted finite epistemic agents never reaches beyond the strongest historically available epistemic justification to directly embrace justification-independent, justification-transcending JTB or J*TB truth condition T. We avoid epistemic hypocrisy in the intended sense by not pretending that knowledge entails the satisfaction of a justification-transcendent truth condition, recognizing instead that in practice all our judgments of truth depend on whether or not we are in possession of a properly qualified justification for believing in the proposition's truth.

Epistemic hypocrisy occurs inevitably in the course of trying to apply the traditional Platonic or Socratic analysis of the concept of knowledge, often in a

Gettier-proof J*TB version of JTB. What we finite fallible epistemic agents actually do in practical situations when our knowledge claims are challenged is to invoke the conclusions of our best efforts at justification. What we think is true is precisely whatever we think is supported by the best justification at our disposal. Finite epistemic agents such as ourselves accordingly do not need to satisfy the JTB or J*TB *truth* condition T independently of satisfying a Best-J justification condition, in an explicitly strengthened Best-JB analysis of the concept of knowledge in competition with JTB and J*TB.

Since we make justification do the real epistemic work in deciding which beliefs to include in or exclude from the category of knowledge, and since it is only honest to admit that our judgment as to what constitutes truth is nothing other than our judgment as to whether a certain proposition meets the demands for best available epistemic justification, since in reality we do not and cannot appeal to truth independently of best justification, we should not continue to ascribe to JTB and J*TB principles that we do not actually follow in practice. We should instead work toward a Best-JB analysis of the concept of knowledge, in which T simply disappears as a justification-independent and justification-transcending condition, and acknowledge the vital role of best justification in our actual epistemological practice.

Advantage 3: Practically Attainable Epistemic Ideal

The concept of a godlike ideal epistemic agent with direct infallible access to the truth adds nothing of epistemic value or utility to our own best efforts to justify our belief in a proposition's truth. We can at best aspire to, as an ideal Kantian regulative principle, while no finite epistemic agent can actually attain, the immediate knowledge of justification-independent and justification-transcending absolute truth that is theoretically available only to an ideal JTB or J*TB epistemic agent.[5]

We are limited in our best judgments of the truth to what historically and cultural-contextually we deem to be the best practically attainable epistemic justification for the propositions we believe ourselves to know, recognizing that we are epistemically fallible, and that any such higher-order knowledge claims are

[5] See Immanuel Kant, *Critique of Pure Reason*, trans. Norman Kemp Smith (New York: St. Martin's Press, 1965), A179-180/B222-223. See A569/B597, where Kant describes the distinction between constitutive and regulative principles in these terms: "Without soaring so high [as to specify the unconditionally necessary qualities of a constitutive principle], we are yet bound to confess that human reason contains not only ideas, but ideals also, which although they do not have, like the Platonic ideas, creative power, yet have practical power (as regulative principles), and form the basis of the possible perfection of certain actions."

themselves defeasible. Best justification, defined for present purposes as justification in a proposition's truth without better countermanding justification in the truth of the proposition's negation, in contrast, is a practically attainable ideal, because it is defined in pragmatic or instrumentalist terms as the maximally practically attainable justification for belief in a proposition's truth. Attainability of the ideal is already built-into the pragmatic concept of best justification understood as Best-J. The best epistemic justification of which epistemic agents are practically capable, like justification generally, is historically and culturally dependent and consequently ontically supervenient especially on the state of science and scientific method at the time of knowledge claim validation or invalidation efforts. As such it is attainable in part by virtue of being built upon a solid but defeasible foundation of epistemic justification that has already been attained.

Advantage 4: Best Justification (Best-J) is Scientific

What counts as best justification defined as Best-J, as already noted, is explained in relation to the current state of scientific development, scrutinized by philosophical criticism and subject to philosophical approval for specific epistemic applications. The pragmatic success of mathematics and of scientific methods involving observation and experimentation that have been developed and refined since ancient times, and with increasing momentum since the seventeenth century and European Enlightenment, appear to offer the best prospects with the greatest potential for pragmatic success in confirming or disconfirming knowledge claims, and hence for the discovery and authentication of genuine knowledge. (See also Advantages 5 and 7 below.)

Advantage 5: Borgesian 'Anthill' Justification Disallowed

There is something disconcerting in the traditional JTB analysis of the concept of knowledge in its slack permission of almost any consideration, any facts remotely related to the truth of a knowledge claim, in principle to count as satisfying JTB justification condition J. This is one way in which the Gettier counterexamples get their hold. We see the problem that results from relying too strongly on truth and not sufficiently emphasizing the justification condition of knowledge dramatized to the point of comic absurdity in the imaginary travels narrated by Jorge Luis Borges in his playful ficcione, *Broadie's Report*:

> The common people say they have the power to transform anyone they please into an ant or a tortoise; one individual who noted my incredulity at this report showed me an anthill, as though that were proof.[6]

If we strengthen justification J condition in JTB to Best-J, compensating for the loss of truth condition T, then we avoid flimsy 'justification' in validating knowledge claims that are manifestly unsupported by the best available justification, despite what the 'common people' in Broadie's report may believe themselves to know. The epistemic scenario Borges describes, exhibiting an anthill, does not represent the best justification for the tribe members' belief that they have the power to turn people into ants. Surely, in fact, there is no best justification in accord with scientific method for this bizarre belief, which ought not to count as knowledge against the background of established knowledge.

If we were to describe Borgesian 'anthill' justification in JTB (or J*TB) Land, we would need to note that Borges' tribe members' belief does not constitute knowledge because it is not true. As we more wisely recognize, in that case, what the tribe believes fails the traditional transcendental truth condition for knowledge, and only additionally and independently may not be very well justified. There is nevertheless some, albeit weak, justification for the truth of the proposition that the tribe is capable of transforming humans into ants, just as there seems to be in the fortune teller example, corroborated in this instance merely by the anthill's existence, perhaps because no anthills had ever been remarked prior to their collective efforts at turning humans into ants. Assuming that tribe members are sincere in their bizarre belief, the anthill then provides for them a reason, albeit a laughably weak one, to believe that they can effect such transformations.

If we were subsesquently to describe Borgesian 'anthill' justification in our preferred Best-JB Land, we might begin by observing that Best-JB analysis implies that knowledge by the tribe's members is lacking because their belief that they can turn people into ants (or into a tortoise, none apparently being ready to hand for similar demonstration purposes as the fictional travelogue is narrated) is not supported by the best justification. This means in turn, as Best-J is defined, that both now and at the time when Borges' visitor is supposed to have encountered this doxastically eccentric tribe, there exists better scientific justification, practically available to others if not to the tribe members themselves, for accepting instead the negation of the belief that the tribe can turn people into ants, for believing instead that the tribe can do no such thing. Common sense and

[6] Jorge Luis Borges, "Broadie's Report" (1970), in *Collected Fictions*, trans. Andrew Hurley (London: Penguin Books, 1998), 405.

experimental science here coincide, as they often gratifyingly do, in ruling out palpably preposterous assertions as failing to constitute knowledge.[7]

Advantage 6: Gettier Counterexamples Solution

If we diagnose Gettier counterexamples to JTB as depending on a *thick* interpretation of what is to count as *justification* in meeting the terms of the traditional JTB analysis of knowledge, then we can avoid Gettier counterexamples without appealing to an ideal epistemic agent's justification-independent or justification-transcending access to a proposition's truth or falsehood by strengthening the justification condition to best justification Best-J.

In a composite Gettier-type counterexample to the traditional JTB concept of knowledge, Smith sees Jones drive an Audi TT every day and park it at his house. Smith concludes from this pattern of observation that Jones owns an Audi. Jones, coincidentally, does in fact own an Audi, kept at a distant location and never driven by Jones, although it is not the rented model Smith sees Jones drive. Smith, accordingly, has JTB, but not knowledge, as the Gettier counterexamples are usually interpreted. The JTB analysis of the concept of knowledge fails precisely because of such Gettier-type scenarios.[8]

---

[7]If we eliminate truth condition T from JTB and J*TB, then we have departed from the classical Platonic-Socratic analysis of the concept of knowledge. The Platonic-Socratic approach by default permits justification of almost any kind or strength to satisfy the justification J requirement in JTB, on the grounds, presumably, that the truth of a true belief need only be supplemented by some kind of reason or warrant in order to constitute knowledge. Lulling knowledge seekers and claimants into a false sense of security concerning the strength of justification needed for knowledge on the assumption that what is known is after all true, might be called the *JTB Lullaby*. We playfully set this knowledge theory soporific to music, taking the first eight-plus bars of Brahm's lullaby (*das Wiegenlied* — Opus 49, No. 4) for the purpose, and adding the lyrics:

Do not worry much about
strength of just-i-fi-ca-tion.
For as long as you have truth,
you don't need a lot of proof.

The JTB Lullaby does not make sense for pragmatically-contexted finite epistemic agents, but at most only for godlike ideal epistemic agents. JTB, unsurprisingly, originating with Plato, analyzes the concept of knowledge accessible to a godlike ideal epistemic agent who is simply supposed to have justification-transcending direct access to truth as a condition of knowledge. JTB, as a result, and as we have emphasized, is not guaranteed appropriate for far-from-ideal finite epistemic agents such as ourselves.

[8] Another approach to the Gettier problem is offered by Dale Jacquette, "Is Nondefectively Justified True Belief Knowledge?" *Ratio* 9 (1996): 115-127.

If we describe Gettier in Best-J Land, we obtain something like the following picture. Smith does not have best justified Best-J belief that Jones owns an Audi, based merely on seeing Jones repeatedly drive such a car and park it at his home. The best justification for the truth of the proposition that Jones owns an Audi is in fact overturned by better justification for the proposition's negation. That Jones does not own an Audi, and in particular that Jones does not own the Audi that Smith sees Jones drive, is better justified by the rental agreement records for the Audi Smith sees Jones drive, and the lack of any authentic relevant purchase documents. It is *true* all along, we may suppose with Gettier, that Jones owns an (other) Audi, but it is equally true that Smith, on the basis of his meager justification of merely seeing Jones drive an Audi that turns out to be rented, does not on such a slender basis *know* that Jones owns an Audi.

What if the rental papers for the Audi TT are forged?[9] This unlikely but conceivable circumstance is also readily accommodated by the Best-JB analysis. If the documents are forged, then of course they are not the best justification for either believing or disbelieving that Jones owns the Audi. If the papers are counterfeit, then either there exists or fails to exist evidence of their forgery. If the evidence exists, then it belongs to the collective best evidence that is practically available and that ideally needs to be consulted in rendering a verdict on whether or not Jones owns the car. If such evidence absolutely does not exist, and the living memories of all the persons involved in the necessary transactions have somehow been wiped out, as sometimes happens in philosophical parables, so that no one could ever come to know that Jones does not actually own the car he is frequently seen driving, then the epistemic situation reverts to that discussed in Advantage 8 below, involving a special application of the universal ignorance problem.

Advantage 7: Reversals of Knowledge Validations

We attain to best justification Best-J when we avail ourselves of the best science of our day. We determine to the best of our abilities the best practically attainable justification by critically screening and philosophically approving what appear to be pragmatically the most successful methods of science. We turn to scientific explanations, and the observationally and experimentally established empirical truths by which they are discovered and which they in turn support, in order to arrive at a sense of how the best practically available epistemic justification for a

[9] I owe consideration of this problem to Richard Fumerton.

given knowledge claim at a particular time and under particular circumstances should be understood.

We cannot fault a forensic scientist investigating a crime in 1941 for not using evidence of DNA analysis in order to identify a suspect. But we can and should fault another scientist investigating the same kind of crime in 2012 for not doing so. Suppose, then, that a scientist decides in 1941 on the basis of available evidence and techniques at the time that an actually innocent person is guilty of an illegal act. DNA evidence, unavailable in 1941, would have exonerated the accused. The best evidence today justifies our believing that the accused person did not commit the crime. By the standards science has since developed as arbiter of the best practically available epistemic justification, the evidence by which we now judge that the accused was not guilty is itself objectively justifiable as epistemically superior to the evidence by which in the past it was judged that the accused was guilty.

The imaginary scientist's knowledge claim may have been best justified in 1941, at least in the sense that the scientist was best justified in making the knowledge claim then, even if the content of the knowledge claim itself does not turn out to stand the test of time. It may have been believed at the time as a result to be known that the accused committed the crime. We suppose that new evidence that later comes to light implies that the scientist in 1941, satisfying the best practically attainable standards of epistemic justification at the time, did not actually know, and indeed, that we today do not know and should no longer believe, that the accused person actually committed the crime. To speak otherwise is to deny the obvious fact that we sometimes change our beliefs as to whether or not a given knowledge claim is true. When we do, we appeal ideally again to the best justification available at the time of the knowledge claim's evaluation. We do so because in fact we cannot do otherwise, unless we pretend to an occult justification-independent familiarity with absolute justification-transcendent truth, reintroducing a condition we have already discredited. Such judgments of a knowledge claim's truth value, satisfying the proposed analysis of propositional knowledge as Best-JB, as a rule are intuitively reasonable, plausible, and squarely in accord with common sense.

Advantage 8: Solution to Universal Ignorance Problem

What happens when everyone agrees throughout the entire history of the human species, past, present and future, that a knowledge claim $\exists sKsp$ is true when p is actually false? The answer in the Best-JB analysis of the concept of knowledge is that

in that case everyone was wrong, and went to their graves one and all defeasibly but manifestly not best justified in believing that *p* itself and ∃*sKsp* are true.

There must then exist countermanding better justification for the contrary belief that not-*p*. Such justification must include and anyway depend on the corresponding nonexistence of the state of affairs whose existence is proposed by proposition *p*. It is hypothetically available in principle to knowledge seekers, but, as world events happen to transpire, it never occurs to any thinker in human history for consideration, despite being hypothetically timelessly true. The situation, then, is that we are supposing omnisciently and in other ways epistemically fictionally that better justification exists in fact for the negation of a universally believed proposition, even though no human being ever happens to become aware of the recalcitrant facts that make all contrary knowledge claims timelessly false. The objective facts of the world make the beliefs of knowledge seekers in the thought experiment timelessly false, without their ever happening to become aware of a discorrespondence between the contents of their beliefs and the relevant actual state of affairs.

What, then, if there's an evil demon who always misdirects me, even when I supposedly have best evidence? Do I then have knowledge? This is similar to the universal ignorance case and should be treated as such on the more limited scale of a single individual's lifetime. The claimant does not then have knowledge, because, although unaware of it, despite being defeasibly best justified, the knowledge claim in question is assumed to be false. The evil demon does not add anything to or subtract anything from identical epistemic situations that can arise for different reasons without assuming the demon's epistemic deviltry. I may believe that I satisfy the requirements for knowledge, in that case, but in fact I do not. If you are in a position to say that I do not know what I claim to know, then you must also have access to better justification than I do, in which case I do not after all have best justification in the Best-J sense. Otherwise, entertaining the logical possibility that the knowledge claim under attack is not actually best justified from a practical standpoint is theoretically and methodologically unintelligible beyond the pat acknowledgement that all best justification and hence all knowledge within the Best-J model is defeasible, and all cognitive subjects are fallible.

From the dialectical standpoint in which Best-JB is offered as a pragmatic replacement for the godlike ideal epistemic agent JTB or J*TB analysis of the concept of knowledge, thought experiments about a knowledge claim being false cannot be motivated or understood except on the assumption that, even if we are not practically in a position to provide it, there exists better justification for the knowledge claim's negation than our imagined justification for its truth. It is

always possible in principle for such an epistemic turnaround to occur, because the possibility is built into the concept of *defeasible* best justification that I do not know or do not know that I know what I believe myself to know. I may claim to know something, even with best justification in my judgment that my knowledge claim is true, from which it still does not logically follow that I actually know what I claim to know. In fact, or by hypothesis, it can happen that I do not know what I believe and judge myself to know, but falsely believe that I know. The philosophical point in relation to the proposed analysis of the concept of knowledge is that this situation is unintelligible also for the JTB or J*TB proponent in lieu of the assumption that there exists a better justification for the negation of a belief than for the belief itself, even if for circumstantial reasons we can never actually lay hands on it.

If it is asked in conclusion whether Best-JB is supposed to be an analysis of knowledge itself or only of best justified knowledge claims, the answer is that Best-JB analyzes the concept of knowledge, but that its applications, whenever we get down to individual cases, can only address the justification status of particular knowledge claims.

6. Objection to Best-JB: The Pregnancy Test-Kit Counterexample

We next address a criticism of Best-JB based on a thought experiment involving a pregnancy test. The test instructions are followed and the result says +, but in fact – is true instead, and there is at the time in some sense supposedly no better available justification for the contrary judgment that –. Does the kit user know that she is pregnant? Presumably not. If, however, we are supposing that – rather than + is true, then we must also be supposing that there exists at the time better evidence for + as the negation of –.

Such better countermanding evidence need not realistically be practically accessible to the kit user, even if only for such mundane reasons as financial, but could in principle include a doctor's examination, which, with the right equipment and procedures, would undoubtedly constitute better evidence than the kit user's drugstore test. The latter is likely to be neither as sensitive nor as accurate than the best that modern medical science can provide in deciding the question of a woman's pregnancy. Even the old fashioned blood test, hard as it was on the bunny population in those days, is presumably more definitive than an off-the-shelf box of test strips that might work reliably if used properly, but that can easily be mishandled in a number of ways that could invalidate the results, if the chemicals were to be contaminated or go prematurely stale, or the like.

If the fact is that –, then why would a properly manufactured chemical pregnancy test kit instead give the result that +? We must suppose at the same time that there is something defective about the pregnancy test kit, and that therefore, unknown to the kit user, the kit's evidence is not the best justification that the kit user is not actually pregnant. Again, the appearance of counterexample is fostered only by describing a situation in which the requirements for best justification may appear to be but are not actually met, and it is assumed as though from on transcendental high that the justified but not best justified belief is not actually true. If the best justification practically available to the kit user in the logically most narrow sense supports the truth of – rather than +, or + rather than –, then the kit user is anyway in the same narrow sense best justified in judging that she knows that she is pregnant, whereas under the transcendental counterexample assumption she does not know that she is pregnant.[10]

The reply is that if we can abstract from the kit user's actual epistemic situation to stipulate that what she claims to know on the basis of her home pregnancy test is that she is pregnant when she is not, then we ought in all fairness also to be able to abstract from her actual epistemic situation to remark that in that case, if her belief is actually false, then there must exist better justification for the negation of her belief. This information makes her justification *ipso facto* something significantly less than best, whether she knows it or not (defective kit, improper application or interpretation, etc.). Otherwise, the imagined counterexample is unintelligible, even on a JTB or J*TB analysis of the concept of knowledge. In general, these challenges pose no worse problems for Best-JB than they do for JTB or J*TB.[11]

[10] The home pregnancy test kit problem was suggested to me by Andrew Moon.

[11] Versions of this essay were presented under the same title at the Episteme Conference on "Justification Revisited," Université de Genève, Geneva, Switzerland, March 25-27, 2010; as "Against Epistemic Hypocrisy," at the Copenhagen-Lund Workshops in Social Epistemology, Copenhagen University, Copenhagen, Denmark, November 25, 2010; and as "Knowledge Without Truth," at the Philosophy Colloquium, Institut de philosophie, Faculté des lettres et sciences humaines, Université de Neuchâtel, Switzerland, November 16, 2010. My thanks to many participants at these venues for offering useful comments and criticisms.

SCEPTICAL THOUGHTS ON PHILOSOPHICAL EXPERTISE

Jimmy Alfonso LICON

ABSTRACT: My topic is two-fold: a reductive account of expertise as an epistemic phenomenon, and applying the reductive account to the question of whether or not philosophers enjoy expertise. I conclude, on the basis of the reductive account, that even though philosophers enjoy something akin to second-order expertise (i.e. they are often experts on the positions of other philosophers, current trends in the philosophical literature, the history of philosophy, conceptual analysis and so on), they nevertheless lack first-order philosophical expertise (i.e. expertise on philosophical positions themselves such as the nature of mind, causality, normativity and so forth). Throughout the paper, I respond to potential objections.

KEYWORDS: expertise, philosophical methodology, reliability, dissensus

1

Although there is a great of talk about expertise in the epistemology of disagreement,[1] there is little, if any, discussion on the nature of expertise; e.g. discussion of the ascription conditions of expertise to others. In this paper, I defend a reductive analysis of expertise as the most plausible account available, along with its (sceptical) implications for philosophical practice – I argue that although philosophers might enjoy something akin to expertise of some kind or other (e.g. they are good at critical thinking, fine-grained distinctions, issues in the history of philosophy and so forth), it is implausible, at least based on the available evidence, that they have the same kind of expertise as scientists, for instance. Throughout the paper, I defend my account of expertise, and its consequences, against potential objections.

2

Surely, we believe that there are individuals who are rightly considered experts; e.g. individuals from engineers, scientists and mathematicians are experts in their

[1] See: John Beatty, "Masking Disagreement among Experts," *Episteme* 3, 1-2 (2006): 52-67; Michael Bergmann, "Rational Disagreement after Full Disclosure," *Episteme* 6, 3 (2009): 336-353; Earl Conee, "Peerage," *Episteme* 6, 3 (2009): 313-323; Axel Gelfert, "Who is an Epistemic Peer?" *Logos and Episteme* 2, 4 (2011): 507-514; Alvin Goldman, "Experts: Which Ones Should You Trust?" *Philosophy and Phenomenological Research* 63,1 (2011): 85-110.

respective fields. Put differently, we think that there are some people, as opposed to others, who enjoy a privileged epistemic position with regard to a particular body of knowledge. This raises two related issues: (a) specifying the plausible necessary conditions, by and large, of expertise and a method for sorting experts and non-experts and (b) if we can know whether we accurately ascribed expertise to someone.

In the next couple of sections, I discuss each of these issues in turn.

It is tempting to think that expertise is cashed out exclusively in terms of privileged access to knowing-that;[2] i.e. someone is an expert just in case they know more about their respective subject than most others in their peer group. For instance, someone with a photographic memory, might be able to read a series of books on medicine, chemistry, biology and so forth and, as a result, gain a great deal of apparent expertise necessary to serve as a medical doctor. Under the proposed account of expertise, such an individual would count as an expert.

With the example in mind, consider the following (preliminary) account of expertise:

> (1a) S is an expert with regard to X just in case, *ceteris paribus*, S is more likely to have true, justified beliefs with regard to X than the majority of her peer group.

The account of expertise specified by (1a) can be broken into two separate components: (1a') S is an expert with regard to X just in case X tends to form reliable opinions with regard to X; call this *the reliability condition*. Next, (1a'') S is an expert with regard to X just in case X tends, proportionately speaking, to be more reliable in her beliefs with respect to X than the majority of those in her peer group; call this *the scarcity condition*.

Although the reliability and scarcity conditions are plausible necessary conditions of expertise – that is, we tend to think that experts should be reliable in their opinions and in the minority epistemically speaking, i.e. for the most part, if everyone is nearly as good at knowing X, then knowing X is not a sign of expertise (e.g. generally speaking, we are not experts on what it is a like to be in pain) – there is something missing from (1a) in its characterization of what is it to be an

[2] See: Jeremy Fantl, "Knowing-How and Knowing-That," *Philosophy Compass* 3, 3 (2008): 451-470; Stephen Hetherington, "Knowing-That, Knowing-How, and Knowing Philosophically," *Grazer Philosophische Studien* 77, 1 (2008): 307-324; Paul Snowdon, "Knowing how and knowing that: A distinction reconsidered," *Proceedings of the Aristotelian Society* 104, 1 (2003): 1-29; and, of course: Jason Stanley and Timothy Williamson, "Knowing How," *Journal of Philosophy* 98, 8 (2001): 411-444. For the sake of this paper, I respect the distinction between knowing-how and knowing-that.

expert. Perhaps the following example will better illuminate the relevant intuition.

Suppose that Bob studied everything about medicine that he could get his hands on; he read extensively on biochemistry, human anatomy, pharmacology and other relevant scientific topics. As a result of his extensive learning, the government hired Bob as part of their medic training program in the armed forces. For the first couple of months, Bob was excellent: as medical situations arose, he was easily, and quickly, able to identify them and explain all of the relevant facts of the case. Unfortunately, as time passed, Bob encountered a greater number of situations without knowing how to deal with them; especially if the cases he tackled were not explicitly mentioned in the medical texts he consulted.

Surely, in light of the complications encountered by Bob, he is not really an expert on medicine, even if he enjoyed some components of expertise such as exhaustive knowledge-that. The Bob thought experiment suggests the following, *improved*, account of expertise:

> (1b) S is an expert with regard to X just in case, *ceteris paribus*, (i) S is more likely to have true, justified beliefs with regard to X than most of those in her peer group and (ii) S is more likely to know the relevant heuristics and methods for applying his true, justified beliefs with regard to X.

The addition to our account of expertise is based on the following intuition: an expert with regard to X should be able to, at least to a greater degree than his non-expert peers, improvise solutions to novel problems related to X. Surely, experts should be better equipped, than their non-expert peers, to handle novel problems relating to their field of expertise. Put differently, an expert in X should have methods and heuristics to able to apply to novel situations that arise related to her area of expertise. Put differently: an expert with respect to X should have greater knowledge-that and knowledge-how than most of her peers, *ceteris paribus*.

Propositions (1a) and (1b) are both reductive and fallibilist accounts of expertise. They hold that expertise with regard to X is just a kind of epistemic privilege with respect to X; thus, there is only a difference of degree, rather than kind, between an expert and novice. I assume something like a reliabilist account of justification[3]; that is, a doxastic state is justified if it was produced by a reliable

[3] Alvin I. Goldman, "Naturalistic Epistemology and Reliabilism," *Midwest Studies in Philosophy* 19, 1 (1994): 301-320; John Greco, "Agent Reliabilism," *Philosophical Perspectives* 13 (1999): 273-296; Jarrett Leplin, "In Defense of Reliabilism," *Philosophical Studies* 134, 1 (2007): 31-42;

process of some kind – a process is reliable just in case its produces a far greater number of true, rather than false, outputs over a sufficient period of time. Thus, if something is reliable, then there is conceptual room for its fallibility. Reliability allows for false outputs; any viable account of expertise must allow that experts are highly fallible.

For our purposes, I take it that the components of proposition (1b) are necessary conditions of expertise, or minimally, plausible candidates for necessary conditions. Proposition (1b) is made up of three components:

The reliability condition	S is an expert in X just in case S is a reliable source of outputs relating to X.
The scarcity condition	Within their community, experts with regard to X should tend to be in the minority.
The heuristic condition	S enjoys expertise with regard to X just in case S knows, more so than those who are not experts, how to apply what knows she about X to novel situations.

With a basic account of expertise in hand, I apply proposition (1b) to some examples.

Consider the following examples. Suppose that Mary is an economist who specializes in stock market trends. Whenever there is ever so slight a shift in stock prices, Mary tends to have already made the relevant predictions. We might suppose, for the sake of the example, that Mary has developed an advanced mathematical model that does the majority of the nitty-gritty calculations and predictions, even though Mary is responsible for its design and execution. Surely, even in the face of a few failed predictions from time to time, we would ascribe expertise in stock price prediction to Mary; she clearly meets all three of the plausible necessary conditions of expertise – assuming, of course, that her skill set is comparatively rare in her peer group.

Now consider Sam, a theoretical physicist working on a predictive model far beyond anything anticipated by his colleagues. As it turns out, long after Sam's death, his predictive model is shown to completely accurate, minus a few details, to the extent that it anticipated theoretical problems that would not have been intelligible to other, equally well-trained physicists, who were contemporaries of Sam. Surely, in this case, we would ascribe expertise in predicting future trends in physics to Sam posthumously, even relative to his expert colleagues.

Michael Levin, "You Can Always Count on Reliabilism," *Philosophy and Phenomenological Research* 57, 3 (1997): 607-617.

3

Now that we have a rough idea of the conditions for expertise, I want to focus on the conditions we tend to follow ascribing expertise to others. Although we lack unmediated access to whether an individual enjoys truth-tracking doxastic practices – for instances, I cannot just *see* that someone has true beliefs; to see why consider the following argument: we value justification instrumentally because it is a reliable guide to discerning truth-tracking propositional attitudes[4] from those that are not – nonetheless, there are indicators of expertise that we right depend on. I will discuss a couple indicators: *institutional trust and innovation.*

To begin with, consider institutional trust. We recognize colleges, universities, job-training programs, apprentice-ships and so forth as institutions that, if employed properly, confer expertise on those who complete the programs. This is one reason we place epistemic value, to a greater or lesser extent, on graduating from an institutional of higher learning. At least in many cases, such is a good indication that one is closer to expertise with regard to their field of study than someone else without that sort of training, *ceteris paribus*. Notice, however, that expertise need not be conferred by a place of higher learning; being an apprentice to a master is another reliable indicator of expertise. That is, having successfully studied a particular practice under the direction of an established master confers indicators of expertise in the relevant field.[5]

Now, consider innovation – that is, meeting a set of specified success conditions that are rarely met. The ability to meet success conditions with regard to a particular area, reliably and on a regular basis, is a good indicator of expertise. Suppose that someone consistently uses a handheld device to locate gas and water pipes buried deep in the ground, where just about everyone else fails in this task; this is a good indication that the individual enjoys expertise with regard to locating gas and water pipes. Or, consider another example: someone who is skilled at carpentry enjoys the necessary expertise to build houses; e.g. those who have built houses for years surely enjoy home construction expertise.

With the conditions for expertise, and ascribing expertise in place, I want to apply the account of expertise we've been developing to the supposed expertise of philosophers.

---

[4] Laurence BonJour, *The Structure of Empirical Knowledge* (Cambridge: Harvard University Press, 1985), 8.

[5] I am not going to discuss how the institutions that confer expertise indicators gain their epistemic credentials for fear that the discussion will lead to the problem of easy knowledge. See: Stewart Cohen, "Basic Knowledge and the Problem of Easy Knowledge," *Philosophy and Phenomenological Research* 65, 2 (2002): 309-329.

4

Generally, we do not think that professional philosophers enjoy the same kind of expertise as scientists. Obviously, there is a sense in which professional philosophers enjoy expertise; that is, they are experts in that they have the ability to explain the view of a particular philosopher on such-and-such issue. For instance, there are experts who are well-qualified to talk about Hobbes' view of human freedom as it relates to his view of political sovereignty; or, there are philosophers who are more than qualified to explain the differences among the various Humean forks. Let us call this, second-order expertise; that is, it is expertise about the philosophical opinions *themselves* rather than expertise about philosophical matters.

There is a major factor that motivates our scepticism directed at the supposed first-order expertise of philosophers: the contentious methods they employ. To clarify the point, consider a passage from Earlenbaugh and Molyneux:

> […] one is not typically inclined to believe P on the basis of someone else intuiting that P. In this way, intuitions come apart from the standard basic evidential states, for no standard basic evidential state is subject sensitive in this way. One is willing to treat what other people seemed to see, what they seemed to hear, and what they seem to remember, as evidence, in the sense that one will base one's beliefs upon it. But one is not willing to base one's philosophical beliefs on the intuitions of another.[6]

For the most part, people do not treat the conceptual machinery and dialectic devices often used by philosophers, e.g. intuitions, thought experiments, fine-grained distinctions etc., as the sort of methods that enjoy evidential momentum; at least compared to perceptual and testimonial evidence, for example.

To clarify the issue, I will pursue it in greater detail.

Why think that philosophers lack first-order expertise? The argument stems from the reliability component of (1b) – if a group of philosophers enjoy the same degree of expertise, more or less, then they should enjoy roughly the same degree of reliability in the philosophical beliefs they hold; reliability is directly tied to its ratio of truth-tracking to false-tracking outputs; a reliable process is one that produces a greater number of true, rather than false, outputs over a sufficient period of time.

Consider an example: suppose that we compare calculators to test their accuracy. After a series of calculations, we find that out of twenty calculators, all but one calculator produced the same answer in every case. There is appears to be

[6] Joshua Earlenbaugh and Bernard Molyneux, "Intuitions Are Inclinations to Believe," *Philosophical Studies* 145, 1 (2009): 99.

good evidence for the following conclusion: the nineteen calculators that consistently agreed on the answer are probably reliable, while the single dissenting calculator is probably defective; that is, it is more plausible to suppose that one calculator failed, rather than supposing that nineteen failed, somehow managed to fail exactly the same way, every time. Thus, if something (or someone) is reliable with regard to Y, then on the basis of the truth-tracking nature of reliability, Y-related outputs should generally agree; call this principle, *the nature of reliability.*

Of course, there will be a few instances of disagreement, even among experts. For instance, there might be a paradigm shift in a sub-field of biology, leading to disagreement until a new paradigm is established. However, dissensus, by itself, is not an indicator that individuals lack first-order expertise (i.e. knowing whether some position or other is the case), unless the dissensus is persistent enough; for instance, if physicists were in constant turmoil, over the fundamental areas of their subject, such that they never established a body of fundamental knowledge, we would be much less inclined, than we generally are, to attribute expertise to them.

Now, consider *the problem of dissensus*. If a group of individuals count as experts with respect to X, then they should, generally speaking, have reliable X-related beliefs. It follows that if a group of people are experts with regard to X, then they should, *ceteris paribus*, reach consensus far more often than not – assuming we are talking about the same areas of specialization. But this is not the case with philosophers. Even though they are aware of the arguments, intuitions and thought experiments (including other relevant conceptual machinery) of their opponents, there is little, if any consensus, on just about every topic of philosophical interest; surely, the problem of dissensus, as an indicator of unreliability, is a good reason to think that philosophers lack first-order expertise.

Put differently, although professional philosophers meet the scarcity condition for expertise, they do not meet the heuristic and reliability conditions – or, at least, it does not appear that they do – for ascribing first-order philosophical expertise to them. To the extent that those conditions are necessary conditions of expertise, this is a problem for their supposed first-order expertise.

For instance, Christensen writes:

> If you'd like to make a professional philosopher uncomfortable, try asking for clear examples of our discipline's achievements in settling the questions we study … Of course, the worry is not about any dearth of philosophers with firm opinions on the great questions. It is about how few of these opinions have, over the years, achieved anything like consensus. Lack of consensus might well … be

> taken as evidence that the parties to the dispute lack good reason for confidence in their positions.[7]

Brennan agrees:

> Philosophers disagree immensely in significant ways. Our best philosophers disagree over the doctrines, methods, and even the aims of philosophy. Experts in all fields disagree, but disagreement is more pervasive in philosophy than in most other fields.[8]

Even though professional philosophers are probably better at critical thinking, conceptual analysis and the history of philosophy than their layperson peers, there is no apparent reason to think they have better access to philosophical truths than their layperson, non-philosophical, counterparts. Surely philosophers have second-order expertise in areas such as conceptual analysis, critical thinking and the history of philosophy, but they do not have expertise when it comes to philosophical knowledge itself.

For instance, there is no compelling reason to think that philosophers, compared to their non-philosophical counterparts, are better placed to know whether or not there is a God, or if freedom of the will and causal determinism are compatible. Surely, philosophers are especially good at deciding whether or not an argument is any good. But that does not appear to translate into first-order expertise. To give a concrete example, there is *positive* evidence that ethicists do not act more ethically, generally speaking, than their non-ethicist academic counterparts.[9] However, it seems that *ceteris paribus*, if one has studied ethics extensively, and thus enjoys greater ethical knowledge, then they would be more inclined toward ethical behavior.

With this in mind, consider the following, formalized version, of the argument:[10]

> 1. If individuals A and B are experts with regard to X, then, *ceteris paribus*, they are generally reliable with regard to X (from [the reliability component]).

[7] David Christensen, "Disagreement as Evidence: The Epistemology of. Controversy," *Philosophy Compass* 4, 5 (2009): 756.

[8] Jason Brennan, "Scepticism about Philosophy," *Ratio* 23, 1 (2010): 1.

[9] Eric Schwitzgebel, "Do Ethicists Steal More Books?" *Philosophical Psychology* 22, 6 (2009): 711-725; Eric Schwitzgebel and Joshua Rust, "Do Ethicists and Political Philosophers Vote More Often Than Other Professors?" *Review of Philosophy and Psychology* 1, 2 (2010): 189-199.

[10] A similar argument can be formulated on the basis of the heuristic condition.

2. If A and B are generally reliable with regard to X, then, *ceteris paribus*, they should generally reach consensus with regard to X (from [the nature of reliability]).

3. Professional philosophers generally enjoy dissensus on first-order philosophical matters (from [the dissensus problem]).[11]

4. Thus, professional philosophers are not experts on first-order philosophical matters.

I take it that premises (1) and (2) are generally accepted features of expertise and reliability respectively; that is, irrespective of whether one thinks that professional philosophers enjoy first-order expertise, premises (1) and (2) are plausible. I take premise (3) to be a reasonable conclusion, merely based on the disputes between philosophers since the time of Plato. Those who are skeptical of dissensus among philosophers need only take a course in the history of philosophy. Thus, there is *at least some reason* to think that philosophers lack first-order expertise.

5

Finally, there are critics who will object like so: those who claim that philosophers lack first-order expertise undermine their argument. After all, denying that philosophers, by and large, lack first-order expertise is itself a first-order claim: it is the case that philosophers lack first-order expertise on philosophical matters – at a minimum, it appears that we should doubt that I could know that claim is true for the same reasons that we doubt philosopher's first-order expertise. But if there is no first-order philosophical expertise, then anyone, whether or not they have been trained as a philosopher, can contradict this paper with the same degree of epistemic authority as her trained counterparts.

Surely, there are those who accept the expertise of philosophers on first-order philosophical issues; but, this prospect – the claim that there are critics, with as much evidential authority as the professional sceptic who disagree – undermines the authority for the claim that philosophers lack first-order expertise; after all, if there are no first-order philosophical experts, then one person's opinion is just as good as everyone else, at least on philosophical issues. Thus, the claim that philosophers cannot lack (first-order) expertise appears self-defeating.[12]

---

[11] It is possible that are some reliable philosophers, even though the vast majority of their colleagues are not; although this is possible, it is highly implausible. Rather, we should suppose that either almost everyone is more or less reliable, or nobody is reliable.

[12] There are philosophers who appeal to a similar argument against those who are skeptical of the evidential credentials of philosophical *intuitions*. See: George Bealer, "The a priori," in

The objection is right, at least as far as it goes – if I claimed to *know* that philosophers lack first-order expertise, I would be stuck in the same epistemic boat as other philosophers. However, the objection ultimately misses the point. Throughout the paper, I employ what I call *doxastic motivators*. That is, I appeal to what are hopefully either beliefs, or minimally, inclinations to believe, on the part of my audience, those propositions that are dialectically sympathetic to the point of this paper.

For instance, I take it as granted that many people will be sympathetic to the following claim: if a group of people is reliable with regard to Z, then they should generally reach consensus with regard to Z, *ceteris paribus*. The same degree of plausibility is assumed for the other components of the argument. If there is reason to suppose that philosophers lack first-order expertise, and this paper appears to rely on first-order expertise, then the charitable move would be to interpret the arguments made in this paper in doxastic terms (i.e. non-evidentially).

To conclude, it seems, based on the reductive analysis of expertise, there is good reason to suppose that although philosophers enjoy expertise on second-order philosophical issues, such as conceptual analysis and the history of philosophy, it seems doubtful that they have something akin to first-order expertise, such as whether they have the right position on the nature of freedom, justice, knowledge and so forth.

Blackwell Guide to Epistemology, eds. Ernest Sosa and John Greco (Oxford: Basil Blackwell, 1998), 243-270.

PHILOSOPHICAL PEER DISAGREEMENT

Nicolás LO GUERCIO

ABSTRACT: It has been widely discussed, in recent years, which is the rational doxastic reaction in the face of peer disagreement. But not much has been said about an interesting instance of that debate: *philosophical* peer disagreement. That is precisely what I will be concerned with in this paper. First, I will offer a definition of *philosophical peer* that introduces the idea of an epistemic perspective. The proposed definition allows for a doublé distinction: between Strong and Weak Peers, and between Strong and Weak Disagreements. Based on these distinctions, I will defend that different doxastic reactions are required depending on the type of disagreement at issue. On the one hand, in the face of Weak Disagreement, we should be conciliatory. Cases of Strong disagreement, in turn, shouldn't motívate a doxastic revision. In order to argue for that, some refinements into the notion of Rational Uniqueness will be needed.

KEYWORDS: peer disagreement, epistemic perspective, uniqueness

1. Preliminaries

In recent years peer disagreement has received much attention. There are basically two main approaches to the problem: Conciliationism and Non-Conciliationism. Conciliationists claim that the rational response in the face of peer disagreement demands of the agent to reduce the level of confidence in her own belief. Non-conciliationists, in turn, maintain that the correct attitude is *to stick to one's guns*. Now, one area of interest where peer disagreement may take place is philosophy. Undeniably, disagreement is common between philosophers, and moreover – it is very plausible to suppose – between philosophical peers. The question that arises is: What is rational to do in the face of cases of *philosophical* peer disagreement? One possible approach would be to suppose that whatever you think about peer disagreement in general can be applied to *philosophical* peer disagreement. In the first part of this paper I will argue that that is not the case. Philosophical peer disagreement requires finer-grained notions. In section 2, I'll present a way of treating cases of philosophical peer disagreement that differs from more general approaches to the subject. The main innovation consists in a different treatment of the notion of peer, which allows for more subtle distinctions. In section 3, I will put those distinctions to work to provide an answer to the main question of the paper: when (if at all) should we be conciliatory in cases of *philosophical* peer disagreement. In order to attain that goal, revisiting the notion of Uniqueness will be needed. In section 4, I will draw some conclusions.

2. Philosophical disagreement

The debate over peer disagreement has a particular application to philosophy. Indeed, there is disagreement in philosophy. Moreover, *prima facie*, there are philosophical peers, i.e. philosophers with equivalent philosophical credentials. Now, it may be thought that the conclusions reached in the former debate can be applied straightforwardly to the latter. I disagree, for the case of philosophy presents some peculiarities worth exploring. In this section, I'll be concerned with those peculiarities, introducing some new concepts that will help to shed light on the debate.[1]

One central notion in the debate over peer disagreement is that of *epistemic peer*. The usual characterization in the literature goes along these lines:

> **Epistemic Peer** Two agents are epistemic peers when (1) they have access to (approximately) the same evidence, and (2) they have the same epistemic virtues (they are equally intelligent, attentive, free from bias, etc).

This characterization is insufficient to capture the idea of *philosophical* peer. One of the reasons is that it fails to acknowledge a very important feature of the notion of evidence: being evidence is not a straightforwardly factual property, but a property that a proposition has only relative to some system of epistemic norms, policies, goals, and methodological commitments.[2] To see this, let me put an example. Some philosophers think that intuitions are the main philosophical evidence.[3] Other philosophers believe, to the contrary, that intuitions cannot play an evidential role. Others may believe that only theoretical intuitions can. That's why, as we'll see, it's important to change 'evidence' for 'facts' in the definition above: two philosophers with different opinions about the evidential role of intuitions may nevertheless share the intuitions themselves. So, they acknowledge the same facts (they have the same intuitions), but differ about the theoretical role

[1] I will be working, all along the paper, within the boundaries of certain philosophical assumptions. First of all, I assume that at least some philosophical theses –those we'll be concerned with– are truth-apt. In second place, I will presuppose that at least with respect to some philosophical theses, philosophers hold the doxastic attitude of *belief*, rather than mere acceptance (or some other). Only cases of the former kind will be of our interest. Finally, I will assume that philosophy is an epistemic endeavor, i.e. that the most important within philosophers' goals are the epistemic ones (there are, of course, several epistemic goals: truth, coherence, empirical adequacy, explanatory power, simplicity, among others).

[2] I will say sometimes that evidence is a 'normative' notion, meaning that what things count as evidence doesn't only depend on the facts but also on the epistemic policies one is willing to endorse.

[3] They endorse an epistemic norm along these lines: If you have an undefeated philosophical intuition regarding p, believe p.

those facts play, for one of them takes intuitions to be evidence and the other don't. Thus, it seems that different epistemic perspectives may assing the status of evidence to different facts. Now, it may be thought that the substitution of 'evidence' for 'facts' is not of any help. Even granting that the concept of evidence is itself problematic, and thus should be abandoned, it seems that 'facts' is equally contentious. To see that it's not, let me say something about 'fact.' What I call the 'shared facts' consist in a core of factual (non-relative) propositions regarding which there is a widespread agreement, without any commitment about the evidential role they play. That core of commonly accepted straightforwardly factual propositions constitutes the basic consensus that is needed to have a rational disagreement. In the next section I'll say something more about the notion of fact.

Another source of relativity present whitin the notion of evidence consists in the relative weight that different agents may assing to the same piece of evidence. To continue with the previuos example, philosohers may agree in that intuitions are philosophical evidence, but disagree regarding the relative weight they have, or regarding what kind of intuitions constitute evidence (Intuitions based on semantic competence? Intuitions about counterfactual situations or thought experiments?). For example, I believe that rejecting a strong intuition comes with a high philosophical cost. But I'm sure not everybody agrees. Some philosophers consider that accounting for intuitions is important, but not *that* important. Intuitions are fallible after all, and we are not only in the business of describing and systematizing our intuitions. So I may share the evidence with another agent while endorsing an epistemic policy that confers much more weight to some part of it.

Another worry has to do with determining what we should do with the evidence; some philosophers think, for example, that we should reach a state of reflective equilibrium between all our philosophical intuitions, while others think, to the contrary, that some intuitions are basic. One final reason for acknowledging the implicit relativity present in the definition of philosophical peer is that different perspectives may come with different epistemic goals. Even assuming that all philosophers share the evidence and the idea that truth is the main epistemic goal, they may diverge with respect to the importance or relative weight of other goals (empirical adequacy, predictiveness, explanatory power, coherence, simplicity, etc.).[4] That discrepancy may explain (at least partially) the diversity of philosophical beliefs even between philosophical peers.

[4] There are philosophers that do not even share the idea that truth is the main epistemic goal. See Johnathan Kvanvig, "Truth is not the Primary Epistemic Goal," in *Contemporary Debates in Epistemology,* eds. Matthias Steup and Ernest Sosa (Oxford: Blackwell, 2005), 285-295 and

The things just mentioned I call them an epistemic perspective. An epistemic perspective includes a conception of what facts count as evidence, the relative weight assigned to different kinds of evidence and what epistemic goals are more relevant, i.e. a number of normative and methodological commitments. Different epistemic perspectives may determine differences with respect to all those issues. Thus, in order to make room for the intuitions above mentioned, I propose to introduce the idea of an epistemic perspective into the definition of 'epistemic peer':

> **Strong Epistemic Peer** Two agents are strong epistemic peers when (1) they have approximately the same epistemic virtues, (2) they acknowledge the same facts and (3) their epistemic perspectives are sufficiently alike.
>
> **Weak Epistemic Peer** Two agents are weak epistemic peers when (1) they have approximately the same epistemic virtues, (2) they acknowledge the same facts but (3) their epistemic perspectives relevantly diverge.[5]

Let me examine each part of the new definitions. First of all, the notion of epistemic peer asks for the parties to have the same epistemic virtues. This remains the same as in the orthodox definition of peer. 'Epistemic virtue' is used in a somewhat vague manner to mean that both parties are equally intelligent, free from bias, careful, etc. The second condition, as found in the literature, usually requires the agents to share the evidence. As we saw in the previous paragraph, that condition is misguided, that's why the change of 'evidence' for 'facts' in the new definition. Here, the fundamental idea is that the property of 'being evidence' is not straightforwardly factual, but relative to the endorsement of some system of epistemic norms or policies, some privileged epistemic goals and methodological commitments, etc (what I've called an epistemic perspective).[6] Finally, I introduced the idea of an epistemic perspective. Now, the new definitions allow for finer-grained distinctions regarding philosophical disagreement. We can distinguish between at least two kinds of disagreement:

> **Weak Peer Disagreement** Disagreement between strong epistemic peers.
>
> **Strong Peer Disagreement** Disagreement between weak epistemic peers.

John L. Cruz and Joseph Pollock, "The Chimerical Appeal of Epistemic Externalism," in *The Externalist Challenge*, ed. Richard Schantz (Berlin: Walter De Gruyter, 2004), 125-142.

[5] Here it's important to point out that although I'm presenting the distinction as a clear cut one, difference or similarity in epistemic perspective plausibly comes in degrees. Thus, there is probably many intermediate states between being weak peers and being strong peers. This being said, I will continue talking in the paper, for convenience, as if there were only two possible cases.

[6] Cf. Hartry Field, "Epistemology without Metaphysics," *Philosophical Studies* 143 (2009): 249 - 290.

Disagreement of the first type is between agents with (roughly) the same epistemic perspective. This means that they share (roughly) not only the facts but also the evidence. Plus, they have (approximately) the same epistemic goals (empirical adequacy, coherence, or whatever). In those cases, I'll argue in the next section, the disagreement cannot be explained by a difference in evidence, epistemic virtues or perspective.[7] So the most plausible explanation is that one of them is misapplying their shared perspective. That is, one of them is committing a *performance error* (she does not believe what she should relative to her own perspective). In those cases, the correct epistemic policy is conciliationism. On the other hand, strong peer disagreement takes place only between weak peers, i.e. agents with different epistemic perspectives. In that case, as we'll see, there are several possible explanations for the disagreement, even granting that each party believes exactly what she should relative to her own perspective. As we will see in the next section, in those cases we need not be conciliatory.

3. When should we be conciliatory?

In this section I'll be concerned with the main question of the debate regarding philosophical peer disagreement: when (if at all) should we be conciliatory. The proposal is that we should incur in a doxastic revision only in cases where the epistemic perspective is sufficiently shared (weak disagreements). To see why, it will be very useful to revisit the notion of Uniqueness and its connections with conciliationism. However, that is not to say, as we will see at the end of the section, that strong disagreements are not interesting or fruitful.

It is usually accepted that conciliationism needs to assume the thesis of Uniqueness.[8] The orthodox enunciation of the thesis is along these lines:[9]

> **Uniqueness 1** Given one's total evidence, there is a unique rational doxastic attitude one could take to any proposition.

If the thesis is correct, there cannot be rational disagreement between agents with access to the same body of evidence: necessarily, one of them is

[7] Of course, the disagreement could be explained by a difference in moral or religious perspective, etc. But, in this work, I will focus only on epistemic perspectives.

[8] On the conciliationist side this is defended by Feldman (Richard Feldman, "Reasonable Religious Disagreements," in *Philosophers Without God´s: Meditations on the secular life,* ed. Louise M. Anthony (New York: Oxford University Press, 2011), 204). On the other hand, Kelly argues for the same idea (Thomas Kelly, "Peer Disagreement and Higher Order Evidence," in *Social Epistemology: Essential Readings,* eds. Alvin Goldman and Dennis Whitcomb (New York: Oxford University Press, 2011), 193).

[9] Cf. Roger White, "Epistemic Permissiveness," *Philosophical Perspectives* 19 (2005): 446.

committing a mistake. Conciliationism needs to grant Uniqueness because if it was possible for different agents with access to the same body of evidence to have different doxastic attitudes towards the same proposition, it wouldn't be clear why they should incur in a doxastic revision only in virtue of the disagreement. It would be still an open possibility that both doxastic attitudes were rational. Now, I think Uniqueness also has to be relativized to a perspective. This follows from the previous discussion: being rational has to do with forming beliefs according to one's evidence. But what facts count as evidence depends in part on the relevant epistemic perspective. Thus, what counts as rational depends also on the epistemic perspective. The correct formulation of Uniqueness should be, then, relativized to a perspective:

> **Uniqueness 2** Given one's facts and relative to a given epistemic perspective, there is a unique rational doxastic attitude one could take to any proposition.

As I said, different perspectives may come with different commitments regarding what facts are considered evidence, or what evidence is more relevant. If that's right, it should be clear that different perspectives may allow for different rational doxastic attitudes towards the same proposition, even in the face of the same facts. So Uniqueness stands, but only relative to a perspective. Now we have the elements to ellaborate a bit more the notion of 'fact' presented in the previous section, and see the connections with the other notions I introduced. The main problem with the notion of evidence, as I have already argued, is that is itself normative, and thus relative to an epistemic perspective. This is something that makes previous definitions of peerhood confuse. In contrast, 'Fact', as I employ the term, eludes this problem. The shared 'facts,' I understand, are just a basic core of non-normative (not relative to a system of epistemic norms, or epistemic perspective) propositions regarding which there is a widespread agreement or consensus, or that are widely accepted. That core of factual, non-normative propositions are necessary in order to start a philosophical debate (or any normative debate, one might think). Thus, the shared 'facts' alluded to in the definitions of Peerhood and Uniqueness above, are only those things about which there is a basic consensus, even across fairly different epistemic perspectives.[10] So, to be sure, what is the connection between the notions of 'Facts,' 'Evidence,' 'Uniqueness' and 'Epistemic Perspective'? The facts, as I already noted, consist in a basic core of non-normative propositions that are commonly accepted, and to

[10] To be sure, the set of straightforwardly factual propositions accepted may also be determined by the endorsement of some system of acceptance procedures. But in any case, I'm assuming that they are very widely shared.

which the epistemic perspective (the normative commitments) is applied. The evidence, on the other hand, is a product of that application. Thus, the notion of 'facts' captures the idea that two agents have some very basic consensus regarding a core of straightforwardly factual propositions, while 'evidence' entails a very much stronger shared epistemic perspective. Finally, the idea to be captured in the new definition of Uniqueness is that, even given some basic facts, there are several beliefs that is rational to form, depending on the held epistemic perspective. But provided some epistemic normative commitments, is no longer relative what beliefs I should form. Similarly, if two agents agree on some basic staff and their epistemic normative commitments are sufficiently alike, any disagreement within the beliefs they formed have to come with some fault on the part of one of them. But if their epistemic perspectives are relevantly different, there is not a unique belief that is rational to form, even given the basic facts on which they agree.

Let me now go back to the main question of the paper: What is rational to do in the face of weak and strong disagreement? Well, in cases of weak disagreement, Uniqueness is in order. Rational disagreement is not possible; hence we need to be conciliatory. But in strong disagreements, there is a difference in epistemic perspective. This means that each party could take a different and yet rational doxastic attitude towards a given proposition. So, rational disagreement is indeed possible, and thus we have no reasons to be conciliatory. It may be the case that in spite of holding incompatible doxastic attitudes towards the same proposition, each agent is being rational, i.e. is forming the belief she should relative to her own perspective.

Then, what are the lessons for philosophical peer disagreement?[11] Well, when faced with weak disagreement, one should be conciliatory. The reason is that the only plausible explanation I have for the diversity of opinions is an error on the part of either me or my rival. But I can't be sure that was her and not me who made the mistake. On the other hand, in a case of strong disagreement, there is no reason to be conciliatory. There are several explanations for the diversity of opinions. First of all, is not at all clear that we share the evidence; but even if we do, we may confer different weight to the same evidence. Second, it may be that we have different epistemic goals. Maybe one of us is interested in empirical adequacy, and the other in giving a prescriptive theory. Now, that would be no problem if we had a way of settling which epistemic perspective is better in an objective manner. But we don't have such a criterion at our disposal. Of course, one could say that the philosophical perspective with a better differential of truth

[11] It could be thought that this kind of proposal can be extended to areas beyond philosophy, such as science. I believe it can, though I will not argue for that in this paper.

over error on the beliefs it produces is the one we should maintain. But, first of all, we have no way of knowing which perspective is better in that sense. And second, that recommendation already presupposes an epistemic perspective.

One final remark. It's important to note that not being conciliatory only means that I shouldn't reduce the level of confidence in my belief just in virtue of disagreement, but not that I shouldn't take into account the other's perspective.[12] It follows from what I said that the second-order evidence provided by peer disagreement motivates a doxastic revision only when my opponent sufficiently shares my epistemic perspective (the facts, the evidence, the relative weight of that evidence, the relevant epistemic goals, etc). But, to be sure, that is not to say that when my opponent doesn't share my epistemic perspective I should discard the disagreement as an uninteresting one. In fact, I think that strong disagreement is a very interesting kind of disagreement. Furthermore, my position is compatible with the following situation: after considering my opponent's perspective I may change my mind. I may judge that her perspective is better for some reason, or realized that my onw perspective has some internal incosistencies. In that case, nevertheless, I wouldn't be changing my mind based on the second-order evidence provided by the disagreement, but based on my assessment of her epistemic perspective.

4. Conclusion

The main purpose of this paper was to bring out the idea of an epistemic perspective into the debate over peer disagreement. This debate, as is currently carried on in the literature, is wrong headed, precisely because it fails to acknowledge that feature. As I argued in the first sections of this work, the idea that some corpus of evidence supports some belief p (or some level of confidence in p), as it stands, has to be rejected. The reason to reject that simplistic picture is that 'evidence' is an implicitly relative notion. That is, what things count as evidence and to which extent they support some belief (or degree of confidence in some belief) depends, at least in part, on an epistemic perspective. This simple point has a direct impact on the debate about peer disagreement. In the first place, it allows for an important and useful distinction between different kinds of peers,

[12] At this point, an objection may be raised. 'To take into account' the other's perspective can only be understood as being open to change my mind. But that is incompatible with my recommendation of being non-conciliatory. The idea is this: I shouldn't change my mind – in cases of strong disagreement – only in virtue of the disagreement. But that doesn't mean that I'm not open to revise my perspective based on a more detailed and sophisticated debate.

a distinction that has been neglected in the debate.[13] On second place, the distinction between different kinds of peers makes room for the acknowledgment of different kinds of disagreement. Thus, it is easier to account for different rational responses in the face of peer disagreements. The result is a much more elastic theory to deal with the problem.

One final remark. The particular case of philosophical peer disagreement is a good starting point, for it seems very plausible that any philosopher assumes, either implicitly or explicitly, the kind of epistemic and methodological commitments I summarize in the paper (regarding what things count as evidence, what weight to assign to that evidence, what epistemic goals are priority, etc). However, although I didn't intend to do it in this paper, I believe that the view previously sketched could be defended in general. I leave for future work the task of spelling out such a proposal.[14]

[13] An exception has to be made here. See Mark Vorobej, "Distant Peers," *Metaphilosophy* 42 (2011): 708-722.

[14] I am truly indebted to Ramiro Caso, Justina Diaz Legaspe, Alfonso Losada and Federico Pailos for their insightful comments on previous versions of this paper. I would like to thank also to the GAF group, especially to the epistemology branch, for their support and very helpful discussion on the general issues.

AFTER UNIVERSAL GRAMMAR: THE ECOLOGICAL TURN IN LINGUISTICS

Noah RODERICK

ABSTRACT: Of all the human sciences, linguistics has had perhaps the most success in pivoting itself towards the physical sciences, particularly in the past fifty years with the dominance of Universal Grammar, which is most closely associated with the work of Noam Chomsky. One of the most important implications of Universal Grammar has been that language production in its most natural and optimal state is organized analytically, and thus shares the same organizational logic of other knowledge systems in Western science, such as the binomial taxonomization of nature and analytic geometry. This essay argues that recent challenges to Universal Grammar represent more than just a theoretical dispute within a single discipline; they threaten to undermine the hegemony of analytical knowledge systems in general. While analytical logic has served Western science well, analogical knowledge systems may be able to address problems that analytical logic cannot, such as ecological crises, the limitations of artificial intelligence, and the problems of complex systems. Instead of studying languages as a means of modeling human thought in general, languages should also be studied and preserved as heteronomous knowledge systems which themselves exist as embodied objects within particular ecologies. Rethinking language as existing on a univocal plane with other ecological objects will provide us with new insight on the ethics and epistemology of analogical knowledge production.

KEYWORDS: Universal Grammar, linguistics, Noam Chomsky, Daniel Everett, ecology, artificial intelligence, taxonomy

1. Introduction

Chomsky's program of Universal Grammar (UG) has had remarkable staying power. It has survived a few major revisions by Chomsky himself,[1] and serious challenges to the program over the years are too numerous to list. After UG, the studies driving knowledge in linguistics would no longer come out of the deserts of the American Southwest or from the Siberian tundra, but rather from computer terminals and libraries in Massachusetts. It is therefore ironic that the most

[1] Chomsky's original Standard Theory was revised by other scholars into the Extended Standard Theory, which Chomsky again reformulated into the Revised Extended Standard Theory in the late 1970's (Martin Edwardes, *The Origins of Grammar: An Anthropological Perspective* (London: Continuum, 2010), 29). Universal Grammar's most recent, stripped down (and therefore, highly defensible) system is the aptly named *Minimalist Program*.

notable challenge to the nativist position (into which UG is incorporated) in recent years has come out of an isolated pocket of the Amazonian rainforest.

Daniel Everett has been generating controversy in the linguistics community since 2005, when he published an article in *Current Anthropology* in which he argued that the language of the Pirahã people provides clear exceptions to what was supposed to be *universal* about UG. Everett claimed that "Pirahã culture severely constrains Pirahã grammar in several ways, producing an array of otherwise inexplicable 'gaps' in Pirahã morphosyntax."[2] Everett has a deep, almost romantic attachment to the Pirahã way of life, and so some of his specific claims about cultural constraints upon the Pirahã grammar and lexicon evoke a visceral response to the grammatical universalist and the multiculturalist alike. His observation that the Pirahã have a spatial-experiential rather than an abstract concept of time[3] reminds one, for instance, of deeply entrenched claims that the Romani language has no future tense or words for time or future because Gypsies are naturally fancy-free.[4,5] But the point on which Everett makes his stand against Chomsky is the argument that Pirahã grammar lacks recursion. Under Chomsky's *Minimalist Program*, recursion is the most visible imprint of UG left in the adult speaker, and it is also the clearest manifestation of the language organ which humans alone possess. Recursion, for one, is part of the same function that allows for a counting system in which postliminary numbers can exist in reference to previous numbers.[6] It is also what allows noun phrases to be embedded into other noun phrases *ad infinitum*. Pirahã, Everett argues, contains neither a counting system nor embedded noun phrases.[7]

There are a few reasons why Everett's challenge to UG has garnered so much more public attention than other challenges over the past few decades. For

---

[2] Daniel Everett, "Cultural Constraints on Grammar and Cognition in Pirahã: Another Look at the Design Features of Human Language," *Current Anthropology* 46, 4 (2005): 622.

[3] Everett, "Cultural Constraints," 631

[4] The persistence of this language ⊃ culture myth about Romani is truly astonishing. The major branches Romani in Europe and the Americas have a most unambiguous term for 'tomorrow' (some variation of *tehara*). And although the future tense is often expressed analytically rather than inflectionally, it is hard to miss even in a superficial study of the language. In this respect, Vlax Romani follows the conventions of other Balkan regional grammars, such as Romanian, Albanian, and Serbian (each of which are part of different language families).

[5] Ian Hancock, "Duty and Beauty, Possession and Truth: Lexical Impoverishment as Control," in *Gypsies: An Interdisciplinary Reader*, ed. Diane Tong (New York: Garland, 2008), 121-122.

[6] Under the Minimalist Program, the ability to count is not strictly a part of the language faculty, but arises out of Merge operations, which the language faculty also draws upon to embed phrases (Noam Chomsky, "Of Minds and Language," *Biolinguistics* 1 (2007): 5).

[7] Everett, "Cultural Constraints," 623-627; 628-631.

one, it's a great David and Goliath story. Chomsky stands as a giant at MIT, and Everett developed his argument while teaching at Illinois State University, a mid-tier public school in the American Midwest. This fact alone speaks to Everett's rhetorical acumen, which is at least a match for Chomsky's own. Secondly, while understanding and dissecting the finer points of UG requires at least some training in linguistics, the idea of recursion is fairly easy for the public to grasp. Most importantly, the notion that one towering scientific theory can be toppled by a single inconsistency or superior idea supports a whiggish view of science that the public and the academy alike tend to favor.

My guess is that Everett's argument will not overturn UG, but it might supersede it.[8] Certainly, a study of the cultural constraints on one language would find fellow travelers in contemporary Functional Linguistics without necessarily having to confront the most important tenets of Generative Linguistics. Furthermore, UG advocates might accept Everett's findings but absorb the punch of his argument by retreating to the split between language and communication, arguing that Everett has provided an instance of cultural constraints on narrative structure. I therefore argue that the impact of Everett's claim must be understood within the context of a larger epistemic shift in which grammar-as-system and language-as-object are becoming increasingly bifurcated. In this essay, I shall argue that the epistemology of language is moving away from the path of internalization to the mind that it has been on since the seventeenth century. Instead of being either a cultural or cognitive resource, language is now being objectified as an ecological resource. This shift has profound implications not only for particular scientific projects, such as modeling artificial intelligence, but also for how scientific knowledge is invented, justified, and argued for in language.

[8] Everett's argument that Pirahã lacks recursion would seem pretty easily falsifiable, but subsequent challenges to Everett have shown just how rhetorical the problem is. Uli Sauerland, who has also worked with the Pirahã, believes he has found evidence of embedded clauses within single sentences; however, Everett examined those same sentences and interpreted them as separate sentences (Eugenie Samuel Reich, "War of words over tribal tongue," *Nature* 485, 7397 (2012): 155-156). The argument where utterances end and where sentences begin cannot, of course, be resolved outside of the framework of a formal writing system. And since it is orthodoxy in Linguistics to downplay the importance of writing, and since Pirahã does not have a formal writing system in the first place, I suspect this particular argument about recursion will remain unresolved.

2. Grammar as a Cultural Exoskeleton

Everett claims that the Pirahã express time grammatically as being either "in experience" or "out of experience."[9] For example:

> Pirahã's (*sic*) excitement at seeing a canoe go around a river bend is hard to describe; they see this almost as traveling into another dimension. It is interesting, in light of the postulated cultural constraint on grammar, that there is an important Pirahã term and cultural value for crossing the border between experience and nonexperience.[10]

And the few words the Pirahã have for time are mostly approximations to specific objects and events. For instance, night (*ahoái*) translates to "be at fire," and specific times of the day can be marked by either "low water" or "high water."[11] Such fundamental connections between environment, language, and the way speakers inhabit the world are consistent with the kind of linguistic relativity associated with the Sapir-Whorf Hypothesis which dominated the study of language in the first part of the twentieth century. The connection between experience and language genesis was first assumed by seventeenth century scholars of General Grammar; however, the connection between cultural experience and particular linguistic variation was formalized by Wilhelm von Humboldt in the early nineteenth century, though von Humboldt was more interested in how certain classes of grammar differently shape rational modalities than he was in the specific connections between physical environment, language, and culture. It was only in the wake of Darwinian science that scholars, such as Benjamin Lee Whorf, had epistemic permission to systematize the environment-language-culture effect. Unfortunately, linguistic relativitism became shorthand for the idea that Eskimos have a whole range of words for snow.[12] However, the most important and controversial arguments of linguistic relativity are about the connection between environment and grammatical categories, specifically the descriptive-taxonomic binary.[13] But since the Universal Grammar revolution,

[9] Everett, "Cultural Constraints," 631.

[10] Everett, "Cultural Constraints," 632.

[11] Everett, "Cultural Constraints," 631.

[12] After Boas and then Whorf made their initial observations about how many words 'Eskimos' have for snow, the idea took on the status of urban legend for the public and an object of ridicule for linguists. In fact, as Harrison points out (K. David Harrison, *When Languages Die: The Extinction of the World's Languages and the Erosion of Human Knowledge* (Cary: Oxford University Press, 2008), the Yupik people do have 99 words for sea ice, which, crucially, help them to identify weather patterns (10).

[13] Benjamin Whorf, "Grammatical Categories," *Language* 21, 1 (1945): 1-2.

linguistic relativitism has been, until recently, largely relegated to theories about metaphor and cognition in socio- and cognitive linguistics.

There are really two reasons why a neo-relativist argument such as Everett's might supersede UG. Firstly, while UG has been an enormously productive theory for linguistics and the computational sciences, it still faces an empirical hurdle which its own methods appear not to be able to jump: the question of origins. Chomsky maintains that UG happened about 50,000 years ago as a single mutation, and as a completely separate event from communication.[14] While others, such as Bickerton, argue for a more gradualist approach in which communication and grammar evolved dialectically in a series of events.[15] I cannot in the space of this essay review the important nuances to this argument which other scholars have contributed, but the problem remains that if UG really is both universal and unique to human beings, then there is no adequate way to study its development in non-humans. Stages of grammatical competence can be observed in children, second-language learners, neurologically impaired individuals, and neurological mapping, but such studies only demonstrate the fact of its existence and what it looks like – its origins must still be inferred from theoretical or simulated reconstruction. Secondly, while other means of observing the development of language are promising, they prove problematic for the argument that language capacity appeared as a single mutation in modern humans. For instance, the forkhead box protein P2 (FOXP2 gene) appears to be partially responsible for the motor skills necessary for human language.[16] However, Krause et al.,[17] working from remains at the El Sidrón site in Spain have detected two amino acid substitutions in FOXP2 in Neanderthals which were previously thought only to be present in modern humans.[18] While few are seriously suggesting that this particular mutation in FOXP2 constitutes the nebulous

[14] Chomsky, "Of Minds," 24.

[15] Edwardes, *The Origins of Grammar*, 34.

[16] Wolfgang Enard, Molly Przeworski, Simon E. Fisher, Cecilia S. L. Lai, Victor Wiebe, Takashi Kitano, Anthony P. Monaco, and Svante Pääbo, "Molecular evolution of FOXP2, a gene involved in speech and language," *Nature* 418 (2002): 869-872.

[17] Johannes Krause, Carles Lalueza-Fox, Ludovic Orlando, Wolfgang Enard, Richard E. Green, Hernán A. Burbano, Jean-Jacques Hublin, Catherine Hänni, Javier Fortea, Marco de la Rasilla, Jaume Bertranpetit, Antonio Rosas, and Svante Pääbo, "The Derived FOXP2 Variant of Modern. Humans Was Shared with Neandertals," *Current Biology* 17 (2007): 1908–12.

[18] Antonio Rosas, Almudena Estalrricha, Antonio García-Taberneroa, Markus Bastira, Samuel García-Vargasa, Andrea Sánchez-Meseguera, Rosa Huguetb, Carles Lalueza-Foxc, Ángel Peña-Meliánd, Elena F. Kraniotia, David Santamaríae, Marco de la Rasillae, and Javier Forteae, "Les Néandertaliens d'El Sidrón (Asturies, Espagne): Actualisation d'un nouvel échantillon," *L'anthropologie* 116 (2012): 65.

'language organ' of UG, it does blur the divisions between communication in non-modern humans and language in modern humans. And it further indicates that a program of study which takes language a uniquely human trait as its *sine quo non* may find itself increasingly isolated from the activities of other sciences.

The corollary argument to Everett's refutation of UG vis-à-vis Pirahã is that endangered languages must be preserved, not just because it would be a pity to see them go or because of postcolonial regret, but because languages are discrete systems bearing special knowledge about the environments from which they emerge, both in their grammars as well as in their lexicons. Instead of a cognitive resource out of which we can better understand ourselves as humans and productively circumscribe computational languages, language is an object of ecology whose genesis, behavior, and fate are absolutely tied to other ecological objects. The shared fate of biodiversity and linguistic diversity is indeed quantifiable.[19] Furthermore, linguistic diversity is now being measured in the same terms of genetic diversity by which biodiversity is also measured. For example, the National Geographic's *Enduring Voices Project* identifies several "language hotspots," which are measured by combining the concentration of phylogenetically distinct languages, the number of speakers relative to age, and the level of existing documentation on those languages.[20]

From a UG perspective, endangered languages are interesting variants of the same human language, and the value of studying them would be to analyze them according to a pre-existing schema of linguistic parameters. Such studies operate well within the Kuhnian description of *normal science*, in which the dominant theory acts as a cup into which continuous successions of data are poured – though it is impossible to imagine the cup overflowing. By contrast, the neo-relativist view of language is characterized by singularity and excess. Here, grammar exceeds the boundaries of the mind. Instead of being a finite structure that can combine and recombine linguistic elements to express an infinite amount of concepts, neo-relativist grammar follows from the infinite realm of experience. Grammar, in other words, is a cultural exoskeleton.

Thus, included in the epistemology of neo-relativistic grammar are heterogeneous knowledge systems, such as folk taxonomies. Folk taxonomies are not merely a collection of metaphors or prepositional systems; they are complex interplays of grammatical genders and embodied temporalities. In contrast to the binomial taxonomic system used by Western science, folk taxonomies contain a

[19] Nicholas Ostler, "A Loss for Words," *Foreign Policy* 139 (2003): 30.

[20] "Disappearing Languages," *Enduring Voices Project.* National Geographic Society, n.d. http://travel.nationalgeographic.com/travel/enduring-voices (Accessed May 7, 2012).

low level of analytic engagement and a high level of analogy. Western binomial taxonomy has a high level of analytic engagement because the identity of each species is taken from an abstract and regular system of units (kingdom, phylum, class, etc.), not unlike the way in which analytic geometry applies imaginary straight lines to measure curves. The analytics of binomial taxonomy are also similar to the analytical engagement in UG, which will be discussed later on in depth, wherein sentences are divided up into micro-phrases, each of which becomes a subset of the phrase above it. Folk taxonomies, I would argue, are often characterized by low analytical engagement because the phenomena they describe need not make sense within an abstract and regular framework.

Take, for instance, the Golden spot hogfish. Its binomial classification is *Bodianus perditio*. The species is *perditio*; the genus is *Bodianus*; its family is *Labridae*, and so on. The fish's name for West Nggela speakers of the Solomon Islands, however, is *Roso taranggua*.[21] *Roso* means "young coconut with soft meat," and *taranggau* is the name for a "fish-eating bird of prey."[22] Thus, "The name refers to the soft flesh of these fishes, which may also be the favorite prey item for the taranggau."[23] Latin names in binomial classification may be similarly evocative (just think of the Grizzly bear's *Ursus arctos horribilis*), but the West Nggela name for the Golden spot hogfish is not derived from the combination of regular units. It is classified according to the speakers' collective experience of it. And embedded within that experience are analogies to other experiences, such as that of eating a young coconut. Such experiential classifications may or may not say much about genetic relations, but they may contain vital information about breeding, behavior, edibility, and environment which only people who have interacted with those animals, plants or things for generations can access.

The distinction I make regarding analytical/analogical levels in different knowledge systems, however, should not be seen in terms of a Western/non-Western binary. Analytic/analogical levels amongst knowledge systems are always relative. For example, even within binomial Western taxonomy, such distinctions exist which biologists have not yet fully resolved. Phenetic taxonomy, which is the true descendant of the Linnaean system, relies heavily on similarities in physical characteristics and ecological niches, whereas cladistic taxonomy requires that species be clustered as monophyletic groups (i.e. groups that descend from a common ancestor). Thus, relative to cladistic taxonomy, phenetic taxonomy contains a high level of analogical engagement. The emergence of DNA

[21] Harrison, *When Languages Die*, 43.
[22] Harrison, *When Languages Die*, 43.
[23] Harrison, *When Languages Die*, 43.

sequencing technologies is largely responsible for the dominance of cladistic taxonomy over the past fifty years, and such developments can give the impression that as observational technologies proliferate, science becomes increasingly analytical. That is to say that what we know about the world becomes ever more encoded into discrete and regular systems of order. Yet, the social reality of science, or that which we determine we *need to know* about the world, does not follow such a neat trajectory. Laws regarding the protection of species may rise and fall according to the genetic distance that can be placed between one type of animal and another. Zimmer reports, for example, that according to genetic analysis, the red wolf (*Canis rufus*) in the southeastern United States, which is currently protected, cannot be considered a separate species from the wolves in Canada and the northeastern United States (*C. lycaon*).[24] In fact, because of the enormous amount of interbreeding with coyotes, red wolves might not even be rightfully called 'wolves' at all. However, with their physical traits and the position they occupy in their local ecosystem, red wolves could be considered an *ecotype*, which does not conform to binomial taxonomy, but which may be more relevant information if we are trying to preserve biodiversity within an ecosystem instead within individual genetic pools. Furthermore, organizing flora and fauna according to ecotype lends itself better to the kind of local knowledge generated by experiential, analogical systems such as folk taxonomies. Thus, the study of endangered languages is one crucial place where ecology as a political exigency meets ecology as an epistemological mode.

3. Out of the Mind and into History

Despite the connection Chomsky's one-time debating partner, Michel Foucault, made between the emergence of modern grammatical study and the taxonomization of nature, a UG advocate might object to any real connection between the two on the grounds that human interactions in nature are of a completely different substance from that which undergirds the faculty of language, which emerged as a single genetic mutation.[25] And although Chomsky explicitly rejects mind/body dualism,[26] there still exists in UG a Cartesian problem regarding the difference nature and culture. Chomsky's earlier work on syntax restored Descartes' distinction between Private and Public language (under the

[24] Carl Zimmer, "What is a Species?" *Scientific American* 298, 6 (2008): 72.

[25] Chomsky, "Of Minds," 22-23.

[26] Noam Chomsky, "Language, Politics, and Composition: A Conversation with Noam Chomsky," in *(Inter)views: Cross-Disciplinary Perspectives on Rhetoric and Literacy*, eds. Gary A. Olson and Irene Gale (Carbondale: Southern Illinois University Press, 1991): 77.

guise of competence and performance) which philosophers such as Wittgenstein had since dissolved. For Descartes, Private Language consisted of representations of mental states, which were transcendent to knowledge of those states as external objects. Something similar happens in Chomsky's famous construction, "Colorless green ideas sleep furiously," which is meant to demonstrate that a sentence can be meaningful on a syntactic level, while nearly void of content on a semantic level.[27] The native English speaker need not objectify the sentence in metagrammatical terms (for instance, describing the verb phrase's relationship to the noun phrase) to know that the sentence is correct. Here we get into the murky territory of a priori synthetic knowledge. Thus, I need to intuit that 'colorless' and 'green' form an adjectival phrase connected to the noun 'ideas,' which form a noun phrase, and that 'furiously,' is an adverb connected to the verb, 'sleep,' which in turn exists in a necessary order with the noun phrase. Although, for instance, 'sleep' can be represented as a noun as well as a verb, within this context, it must be intuited as a verb, just as 'ideas' must be intuited as a noun, and so on. Thus, we know the qualities of the construct's constituents a priori, just as if we know what the numbers 5 and 7 represent, we can combine them into a true statement: 5+7=12.[28] Knowledge of the construct is synthetic because we know the sentence is true as a combination of the noun phrase and the verb phrase in a given order.

Chomsky seems to have doubled-down the aprioristic claims for language, indeed arguing that much of what is traditionally thought of as semantics is really just syntax.[29] Thus, the fact that there can be a verb, 'sleep,' and a noun, 'sleep,' arises from the possibility of a sentence like "Sleep sleeps." Contrary to the ideas of 18th century general grammarians, such as Joseph Priestly, the functional shift between, for instance, sleep and sleeps, is not a process that can be traced to Ø-point derivation from either a referent or an interjection,[30] nor is it a historical constant of grammaticalization as it was for the 19th century philologists,[31] but innate and presupposed under UG. A chimpanzee, for example, can learn the symbol for *sleep*, but cannot derive the relationship between *sleep* and *sleeps*. UG presumes that meaning is ultimately dependent upon neither phonology nor morphology. Chomsky cites the fact that

---

[27] Noam Chomsky, *Syntactic Structures* (Berlin: Mouton de Gruyter, 1957), 15.

[28] Immanuel Kant, *Critique of Pure Reason*, trans. Max Müller (New York: Doubleday, 1966), 137.

[29] Chomsky, "Language, Politics," 93.

[30] Joseph Priestley, *English Grammar: Lectures on the Theory of Language and Universal Grammar* (London: G. Smallfield, Hackney, 1883), 151.

[31] Michel Foucault, *The Order of Things: An Archeology of the Human Sciences* (New York: Vintage, 1994), 252.

> … [E]ven the simplest words in many languages have phonological and morphological elements that are silent. The elements that constitute lexical items find their place in the generative procedures that yield the expression, but cannot be detected in the physical signal. For that reason … it seems … that the language acquired must have the basic properties of an internalized explanatory theory.[32]

What is present syntactically may not be expressed outwardly, at the level of sensorimotor systems. For example, the phrase, "John ate what" is interpreted at the syntactic level as the object, "what John ate what," or separately, "for which thing *x*, John ate the thing *x*."[33] This is an example of the internal Merge function, and its existence presumes that the language of thought produces objects of a different substance from the objects which communication produces. Thus, "[c]omplexity, variety, effects of historical accident, and so on, are overwhelmingly restricted to morphology and phonology, the mapping to the sensorimotor interface."[34] Once again, while the mind/body problem is resolved, the metaphysical split between nature and culture would seem to be left as a remainder.

Yet, research on the borders of syntax and semantics within UG has yielded at least one example in which there exists logical parity between the natural computational structure of the mind and contingent products of culture. In the 1970s, Joseph Greenburg, considered as one of the founders modern linguistic typology, pioneered work in classifying languages according to certain structural elements rather than by historical relationships. Perhaps the most basic way of doing this is to observe the orders which subjects, objects, and verbs take in different languages. For instance, English works in an SVO order: "He read the book." By contrast, Japanese goes by SOV:

Kare-wa hon-o yonde

He-SUBJ book-OBJ read

It is natural to expect that English and Japanese word orders would be radically different, given that historically and geographically, they're about as far apart as two languages can be. However, a closely related language, such as German often uses the SOV order as well:

Er hat das Buch gelesen

He has the book read

[32] Chomsky, "Of Minds," 17.
[33] Chomsky, "Of Minds," 21.
[34] Chomsky, "Of Minds," 22.

Technically, this could be said in English, but of course that would mean that 'read' would be an adverb. And a distantly related, but geographically proximate language, such as Irish, can have a totally different (and much rarer) word order altogether:

Léigh sé an leabhar

Read he the book

Such radical differences in syntax, even amongst genetically related and geographically proximate languages points to the notion that languages might be classified according to more abstract criteria than historical relationships or eco-social conditions (as per Whorfian analysis).

In the early 1980s, Chomsky began to integrate this ahistorical classification into the logic of UG. Known as Principles and Parameters theory, it holds that while there are certain universal principles that constrain language generation in general, there are also parameters that shape and constrain particular linguistic expression. Indeed, in recent years, linguists, such as Mark C. Baker[35] have attempted to create a typology of all known languages based upon a hierarchy of linguistic parameters. According to Principles and Parameters, rather than simply learning vocabulary and word order through whole phrases, children learn to build phrases from the bottom up, using a set of ingredients provided by the linguistic data they receive. For example, in response to a question like "Where did he go?" an English speaker is obliged to answer either with a simple prepositional phrase, such as 'into town,' or with a sentence containing a subject: "He went into town." A child learning Kalderash Romani (Kalderashstika[36]), however, becomes aware that she can respond by saying either

Wov gêlo and'o foro.

He went into(the) town

or simply

Gêlo and'o foro.

[35] Mark C. Baker, *The Atoms of Language: The Mind's Hidden Rules of Grammar* (New York: Basic Books, 2001).

[36] Here, I treat Kalderashstika as a separate language, rather than a dialect of Romani. Romani should be considered a language family (just as English is part of the Germanic family) rather than one language with several dialects.

This is known as the *pro-drop* or *null subject* parameter. It can be said to be 'switched off' in English and 'switched on' in Kalderashstika. But linguistic parameters do not simply compare languages by a series of absent-present distinctions. For instance, the fact that English uses particles and Russian does not is not significant enough to put English and Russian on separate parametric tables. If all known languages were grouped by such small distinctions, there would be quite a bit of clutter. Parametric typology instead works as a binary system, with a regular 'if-then' logic (the kind of which one would find in computer programming language). For example, the parametric difference between English and Kalderashstika involves a couple more degrees of difference than the fact that the latter has the pro-drop parameter switched on and the former does not. A language like Kalderashstika has its pro-drop parameter switched on because it also has what is called the *verb attraction* parameter switched on. Baker defines the verb attraction parameter by the following opposition:

> Tense auxiliaries attract the verb to their position.
>
> or
>
> Verbs attract tense auxiliaries to their position.[37]

The verb attraction parameter thus affects the positions of verbs and adverbs relative to the subject. Thus, in English, it is possible to say 'He strongly agrees.' However, in Kalderashstika, one is obliged to say:

> Wov pristanil zurales.
>
> He agrees strongly.

English has its verb attraction parameter turned off, thereby making it impossible to drop its pronouns. Each parametric binary acts as a subset of another binary opposition (See figure 1).

Notice again that most of the language pairs Baker places on the same branches are historically and geographically unrelated. We get the impression that even as the speakers of those languages develop cultures within various ecological spaces and historical configurations, language itself occupies a separate metaphysical plane.

We find the same binary logic in UG when we move from typology of particular languages to phrase-structure analysis of language in general. Formal linguistics employs several different ways of analyzing syntactic structure, including dependency grammar analysis, which places the predicate verb on top

[37] Baker, *The Atoms of Language*, 132.

of its nodal structure. All other words in the clause relate back to that predicate verb (figure 2).

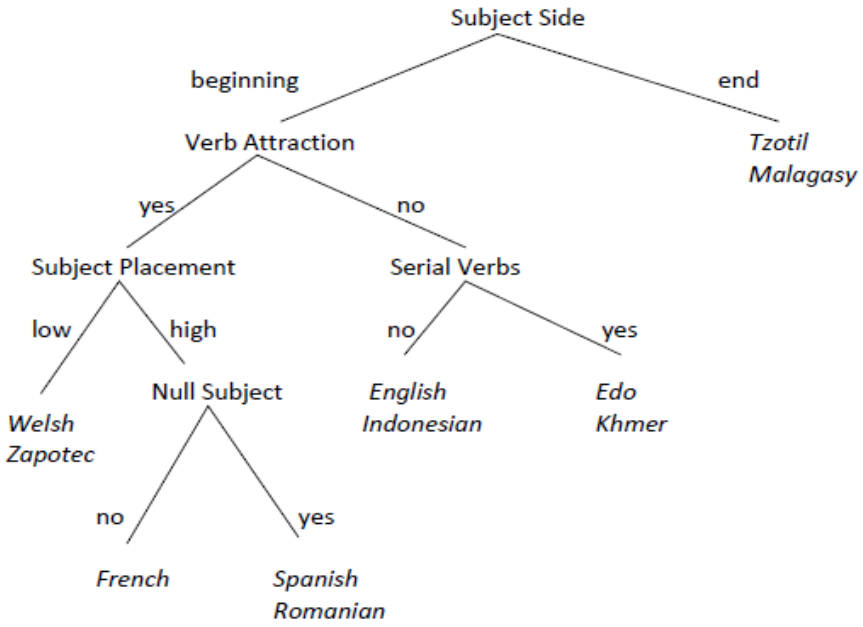

Figure 1: 'The Parametric Hierarchy' adapted from Baker[38]

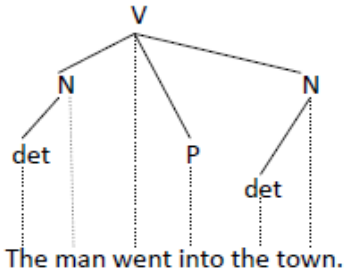

Figure 2: Dependency Grammar Diagram

[38] Baker, *The Atoms of Language*, 183.

Dependency grammar analysis is a somewhat holistic approach in which syntax is driven by lexical elements, taking intentionality as formative. In other words, all other elements in a clause proceed from one focal point, the predicate verb. Chomskyan UG, however, utilizes variants of constituency grammar, specifically the X-Bar approach. In X-Bar theory, clauses and sentences are analyzed in terms of phrasal structure. Functionally, noun phrases, verb phrases, prepositional phrases and so on, are iterations of the same 'X' phrase.[39] X-Bar analysis thus takes on a combinatory and binary structure similar to the parametric typological structure used to classify particular languages:

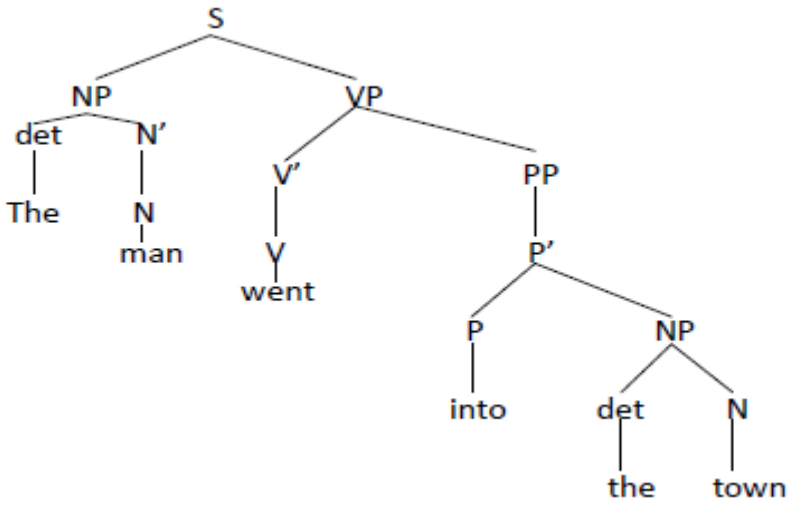

Figure 3: X-Bar Diagram

Despite increasing evidence to the contrary from more sophisticated methods of neural mapping which show that cognitive operations are nodal rather than modal, UG holds to (necessarily so) the idea that mental tasks, such as language are discrete processes. That is not only to say that there is a 'language organ' (even if it is not so easily circumscribable, like the liver), but it is also to say that the interfaces that constitute the module for language create an operation that is unique to the language function. That is why Everett's argument that language is an adaptive tool, similar to the conception and construction of bows and arrows[40] is so abhorrent to UG. Yet, as I have demonstrated with regards to the

[39] Michael Haley and Ronald Lunsford, *Noam Chomsky* (Woodbridge: Twayne, 1994), 144.
[40] Daniel Everett, *Language: The Cultural Tool* (New York: Pantheon Books, 2012), 20.

application of Principles and Parameters to linguistic typology, the same structures that generate individual linguistic expression also govern things in the world: particular languages, such as Polish or Japanese. This is an important step because it crosses an a priori/a posteriori threshold that generative linguistics, by its own theoretical goals, should not cross. The primary divide between generative linguistics and other approaches, such as functional linguistics or relativist linguistics is that the latter describe language as it is actually used, whereas the former describes the existing rules and structures behind the potential for linguistic expression. In functionalist and relativist linguistics, language is already interpellated into the history of material and institutional culture. For instance, in an approach he calls *Sociolinguistic Typology*, Trudgill proposes that grammars move back and forth between simplicity (e.g. increase in morphological transparency) and complexification (e.g. increase in syntagmatic redundancy) through the historical processes of language contact.[41] Grammars, *ceteris paribus*, simplify as a result of contact and complexify in isolation. By contrast, there is no theory of culture in generative linguistics beyond the more abstract implication that humans share discrete instincts for language, morality, and so on. But the parametric theory of how existing languages relate to each other necessarily enters UG into the field of history because the movement of a language from one form to another must pass through certain stages of transformation. Again, if English is to drop its pronouns, then it must also switch its verb attraction parameter on and have its *subject placement* parameter switch to high.[42] That is not to say that parametric typology takes a Hegelian dialectic view of history wherein languages with their *subject side* parameter switched to 'beginning' are moving *towards* dropping their pronouns. But parametric typology nevertheless does offer a theoretical explanation for why certain historical phenomena look the way they do and how they differ from one another.

4. The Mind-Body Split Relocated

It is widely assumed that the difference between nativist and relativist positions is that relativists believe that our worldviews, and thus our particular cultures are shaped by language. For nativists, it is assumed that particular languages do not shape our worldviews. Yet, when it comes to the question of how new knowledge about the world is made, the situation is somewhat reversed. For neo-relativists in particular, language, and thus knowledge, are emergent, following the logic of

[41] Peter Trudgill, *Sociolinguistic Typology: Social Determinants of Linguistic Complexity* (Oxford: Oxford University Press, 2011), 21-22.

[42] Baker, *The Atoms of Language*, 183.

complex adaptive systems. An approach such as Trudgill's, for instance, wherein languages oscillate between grammatical simplification and complexification, follows the logic of thermodynamics. Knowledge creation for nativists, on the other hand, follows the same analytical logic as linguistic creation. Knowledge systems, such as binomial taxonomy, natural language generation, particular languages, and information technologies all begin with a simple set of elements that combine and recombine in a regular way until complexity is achieved. Here, innate reasoning precedes knowledge, and all knowledge is structured like linguistic knowledge.

Again, it is hard to miss the common thread of binary logic that runs through the generation of natural language at the micro-syntactic level to the formation and organization of particular languages to the creation of artificial computational languages. For Chomsky, this is a cause for optimism because it shows that cultural phenomena such as particular languages, as well as advances in human technology, such as information technologies, are reflective of internal and innate reasoning. It's one of the rare places where one might catch a glimpse of the overlap between Chomsky's scientific work and his radical politics. Indeed, his belief that human reason governs linguistic production (and not the other way around) is behind his continued skepticism about the possibility of true artificial intelligence, or at least of our ability to recognize it as such. Here, he cites Alan Turing's own doubts:

> Turing seems to have agreed with Wittgenstein as to the pointlessness of the discussion and debate [over artificial intelligence], until today, over whether machines can (in principle) think, play chess, understand Chinese, do long division, etc., and about how we could 'empirically' establish that they do;…I think Turing's stand was correct.[43]

Chomsky does believe that AI has explanatory uses in terms of modeling thought, but thinks it has no real potential to produce thought in a way we understand it.[44] His stance, like Turing's, is that our notion of autonomous intelligence or thought might change in time, but such a change in definition will not make the phenomenon any more real or accessible to us. The properties that constitute autonomous intelligence would have to be dreamed up as concepts that already correspond to our ideas of intelligence, and so no mind-independent facts could exist to prove or disprove the claim. In other words, we would be making a cultural decision about something which is natural, which, for Chomsky, is unscientific and not worth pursuing. It's a recognizably Cartesian stance.

---

[43] Noam Chomsky, *Language and Thought* (Wakefield: Moyer Bell, 1993), 30.

[44] Chomsky, *Language and Thought*, 92.

It is at this point that the difference between nativists and neo-relativists over how knowledge is produced can be put into the more general context of Giambattista Vico's critique of Descartes' metaphysics. In Descartes' foundationalist philosophy, the goal of all sciences is to be reduced to the analytical truths of mathematics, which are mind-independent: "all the thoughts which can come into the human mind must be arranged in an order like the natural order of numbers."[45] Thus, biology is on its way to becoming chemistry; chemistry is on its way to physics; physics is on its way to geometry; and of course, Descartes himself developed the means to translate geometry to algebra. Vico, in a strikingly Foucauldian fashion, argues instead that "sciences must begin at the point when their subject matter begins."[46] That is to say that knowledge is not already out there waiting to be abstracted from appearance to analytical truth. One cannot, therefore, go in reverse and predict new kinds of knowledge from analytical foundations, just as for Kant, there can be no a posteriori analytic truth. Vico asserts that new kinds of knowledge arise when we encounter things in the world. Further on, new knowledge is created when concepts which formed out of initial contact with things in the world are again encountered in new contexts. Vico's theory of knowledge, in other words, is analogical rather than analytical. In the case of artificial intelligence, then, Vico might argue that our encounter with the concept of artificial intelligence changes our very knowledge of intelligence; we are not just using 'intelligence' as shallow metaphorical cover for something may or may not exist independently, as Chomsky would have it. Beyond that which God himself knows, as far as Vico is concerned, there is no reality underlying our knowledge of something like 'intelligence' outside of our historical development of that knowledge. It's not that Vico can be claimed as an 18th century postmodernist. He still holds that language can mediate collective experiences; it's just that those collective experiences are recorded in language change, which is itself historically and culturally variable.

Underlying Vico's departure from Descartes is the claim that "the mind does not make itself as it gets to know itself, and since it does not make itself, it does not know the genus or mode by which it makes itself."[47] What is more interesting then than to show whether or not it is possible to prove the existence of artificial intelligence is to go back in language to the moments in which the concept of intelligence was abstracted from another concept, to see when and why

[45] René Decartes, *Philosophical Letters*, trans. Anthony Kenny (Oxford: Clarendon Press, 1970), 5.

[46] Giambattista Vico, *New Science*, trans. Anthony Grafton (New York: Penguin Classics, 1999), 9.

[47] Giambattista Vico, *On the Most Ancient Wisdom of the Italians: Unearthed from the Origins of the Latin Language*, trans. Lucia M. Palmer (Ithaca: Cornell University Press, 1988), 52.

it became a subject for a science. This is why philology is key to Vico's method. Not only do abstract concepts (which become subjects for science) become analogized from concrete lexical items, but abstract concepts are also analogized out of grammatical relationships. For example, Vico claims that the Latin *anima* (soul) arises from the concept of air.[48] But since Latin obliges its speakers to make a nominal gender distinction, there also exists *animo* (spirit). Since, it was supposed (and Vico, of course, agrees) that the masculine agent is active and the feminine agent is passive, the more abstract concept of the Will must result from the male *animo* acting on the female *anima*.[49] The spirit acts upon the soul. Furthermore, since it was supposed that the spirit was the agent of neural matter, the soul was the agent of the veins and blood, neural matter must act on the rest of the body.[50] Thus, the systems of atmosphere, gender, and physiology act analogously. For Vico, the evolution of language precedes our knowledge of the physical relationship between the nervous system and the circulatory system.

After spending years traveling around China and studying the development of Chinese technology, Joseph Needham famously posed the question of why European science had so suddenly outpaced Chinese science since the dawn of Modernity. It was, after all, clear that the Chinese had invented so many crucial technologies, such printing and gun powder, well before the Europeans did. His answer was that the modern Europeans managed to create a culture of science around a single set of metaphysical assumptions. The homogeneity of analytical logic allowed disparate sciences to work under the same truth conditions, regardless of their subject matter. Underlying reality in the physical world mirrored the mind's own ordering processes. This is much more efficient than discovering analogies in an infinite world of signs. However, analogy as a mode of knowledge-making is once again coming into prominence with the rise of complex systems sciences. The idea that, for instance, the difference between ant colonies and traffic patterns[51] is a matter of scale rather than substance opens the door for analogies about any number of systems with emergent properties. Such analogies could be used to understand difficult social problems like urban sprawl, but they could also be used to naturalize the inequalities produced by markets in information economies or, as Sagarin et al propose, deal with people deemed

[48] Vico, *New Science*, 313.
[49] Vico, *New Science*, 313.
[50] Vico, *New Science*, 313.
[51] Melanie Mitchell, *Complexity: A Guided Tour* (New York: Oxford University Press, 2009), 3.

'terrorists' in the same way the immune system deals with viruses.[52] Analogical knowledge-making therefore brings with it a whole new set of ethical and political questions which may or may not apply to analytical knowledge-making.

It is here, I argue, that neo-relativist linguistics can make some of its greatest contributions to science and society. With an expanded concept of grammar, languages can be studied as situated objects within a given ecology in order to learn things that analytical knowledge systems cannot. But also, as we expand our notions of intelligence and fellowship to other beings (e.g. animals, artificial intelligence, virtual relationships), the means by which things are classed (grammatical gender) and related (e.g. terms of kinship) may provide insight into the possibilities and problems of making analogies between objects and systems. Indeed, there is a growing area of study known as *morphological computation*, which suggests that advances in artificial intelligence may not come from attempting to deduce and replicate human reason from UG, but by replicating the kind of intelligence that octopuses possess, which is analogical, decentralized and embodied.[53] Studying language as a similarly embodied system may radically alter our conception of human subjectivity, just as the Cartesian Revolution did for the moderns.

A physicist friend of mine once explained to me the difference between the practical applications of Newtonian gravity and those of Einsteinian relativity. He told me somewhat hyperbolically that we got to the moon on Newton but that we need Einstein to get us to the grocery store (referring to GPS gadgets). Something similar might be said about UG and neo-relativist grammar. The major science media outlets are spoiling for a fight between the nativists and the neo-relativists because it makes a great story and it conforms to the still dominant narrative of progressive scientific truth. But the truth is that the competing theories are not going to cancel each other out on the same epistemological plane because, as I have argued, they are working from very different metaphysical boundaries and have very different social agendas. In other words, the different approaches to language will not do each other in; they will do different things.

[52] Rafe Sagarin, Candice Alcorta, Scott Atran, "Decentralize, adapt and cooperate," *Nature* 465 (2010): 293.

[53] Cecilia Laschi, Matteo Cianchetti, Barbara Mazzolai, Laura Margheri, Maurizio Follador, Paolo Dario, "Soft Robot Arm Inspired by the Octopus," *Advanced Robotics* 26, 7 (2012): 709-710.

MENTAL REPRESENTATIONS AND THE DYNAMIC THEORY OF MIND[1]

Cristinel UNGUREANU

ABSTRACT: In this paper I will investigate the possibility of defending the concept of 'mental representation' against certain contemporary critiques. Some authors, like Anthony Chemero, argue that it is possible to explain offline actions with dynamic concepts. Hence, the dynamic discourse preempts the representational one. I doubt that this is a recommendable strategy. A form of representation is necessary, though one which is different from the classical one. Instead of eliminating the concept of representation (as radical dynamicists do) or of splitting cognitive explanation in two separate discourses (as the adepts of the hybrid cognition version do), I consider that a dynamic concept of 'representation' is a better option. In my view, the higher level order resulted from the complex brain-body-environment coupling can be interpreted as being representational in nature. The dynamic paradigm involves a significant change concerning the intentional nature of representational states: the basic forms of representations are not maps of reality implemented as such in the brain, but limit conditions, attractors constraining the cognitive system's evolution in its space state to reach its goals. On a certain threshold of complexity, the system develops stable attractors and attractor landscapes which could be interpreted as standing for something outside the system. This conception offers the advantages of avoiding preemption argument, of unifying the cognitive explanation and, by its interscalar account, offers dynamic tools for building more complex artificial intelligent systems.

KEYWORDS: anti-representationalism, classic cognition, dynamic systems theory, interscalar account, mental representation.

The anti-representationalist challenge

The dynamic theory of mind (or the embodied mind theory – EMT), proposed already in the 90s of the last century by authors as Varela, Thompson, Rosch and many others,[2] is trying to break definitively with the assumptions of the classic

[1] AKNOWLEDGEMENT: This paper is a result of the project "Transnational Network for Integrated Management of Postdoctoral Research in Communicating Sciences. Institutional building (postdoctoral school) and fellowships program (CommScie)" – POSDRU/ 89/1.5/S/63663, financed under the Sectoral Operational Programme Human Resources Development 2007-2013.

[2] See Francisco Varela, Evan Thompson, and Eleanor Rosch, *The Embodied Mind. Cognitive Science and Human Experience* (Cambridge: MIT Press, 1991); Rodney Brooks, "Intelligence without Representation," in *Mind Design II. Philosophy, Psychology, Artificial Intelligence*.

cognition. Its slogan is to bring the mind back into the body and the body back into the world.[3] A central feature of the classic cognition is that thinking does not operate directly in and on the things in the world, but through their representations. Of course, the action takes place in the world, but the contact with the real world takes place first at the moment of perception and second when the behavioral output is produced.[4] These processes are external to the mind; they do not influence its core operations.

Anticipated by Dreyfus,[5] the change in cognitive sciences begins with Rodney Brooks' new bottom-up perspective on building robots. In his view, a system is intelligent when it autonomously copes in real time with the environmental challenges. The ambition of classic computationalism is to build intelligent systems, able to solve complex tasks (to play chess, to recognize linguistic sequences etc.). Brooks, instead, wants to set the coordinates within which the robot develops its own actions, starting with the simplest ones.[6] The robot is connected to the world in a much simpler way, by an ongoing sensing of it. It needs no internal world model. The world is its own model.[7] Perception is direct, not mediated by representations. Its result is not taken by another module in order to build a detailed map of the environment.[8] Perception and action are simultaneous, they form a causal loop.

Revised and enlarged edition, ed. John Haugeland (Cambridge: MIT Press, 1997), 395-420 (originally published in *Aritificial Intelligence* 47 (1991): 139-159); Robert Port and Tim van Gelder, eds., *Mind as Motion. Explorations in the Dynamics of Cognition* (Cambridge: MIT Press, 1995); Andy Clark, *Being There. Putting Brain, Body and World Together Again*. (Cambridge: MIT Press, 1997).

[3] Michael Wheeler, *Reconstructing the Cognitive World. The Next Step* (Cambridge: MIT Press, 2005), 11.

[4] During perception, after the physical input stimulates the sensorial interface of the cognitive system, the visual cortex computes this input in order to produce a three-dimensional representation from the two-dimensional projection of things on the retina; these representations are taken over then by other modules of the cognitive system in order to search for solutions and build action plans. Cf. David Marr, *Vision. A Computational Investigation into the Human Representation and Processing of Visual Information* (New York: Freeman, 1982), 23; Jerry Fodor, *The Modularity of Mind* (Cambridge: MIT Press, 1983), 102-103; Zenon Pylyshyn, *Seeing and Visualizing. It's Not What You Think* (Cambridge: MIT Press, 2003), ch. 2.

[5] Hubert Dreyfus, *What Computers Can't Do* (New York: Harper and Row, 1972); Hubert Dreyfus, *What Computers Still Can't Do* (Cambridge: MIT Press, 1992).

[6] Brooks, "Intelligence without Representation," 410.

[7] Brooks, "Intelligence without Representation," 406.

[8] Brooks, "Intelligence without Representation," 404.

Brooks not only 'weakens' the idea of cognitive processes as internal mental workings by emphasizing the constitutive role of the environment in shaping intelligent behavior, but also adopts an anti-representationalist stance. He shows that the new robots are able to exhibit intelligent and flexible behavior without mental representations. There are no mental symbols functioning as elements of the reasoning process. The robot acts directly on the stimuli, not on their representations. Other research in this field highlights dynamic aspects of infants' gait development,[9] of sensory-motor activity,[10] of phonological system,[11] of language,[12] of perception and action,[13] of limb movement according to a metronome.[14]

Adepts of representational theory assert that cognitive systems posses a set of stored concepts and that their main task consist in extracting information from the perceived stimuli, then comparing it with the stored concepts by deductive operations and, finally, sending an appropriate message to the motor areas in order to perform an action. There are many critiques against this linear process. For example, how does the cognitive system know under what concepts to subsume the features of the current situation?[15] The same problem arises for that representationalist who does not subscribe to inneism. An empiricist must explain how the system succeeds in generalizing the relevant cases, given that there are no absolutely identical situations.[16] Roughly speaking, the representationalists encounter either the problem of relevance or of generalization because the number of the relevant things or of the ways of classifying them is, theoretically,

[9] Esther Thelen and Linda Smith, *A Dynamic Systems Approach to the Development of Cognition and Action* (Cambridge: MIT Press, 1994).

[10] Elliot Saltzman, "Dynamics and Coordinate Systems in Sensorimotor Activity," in *Mind as Motion*, 149-174.

[11] Catherine Browman and Louis Goldstein, "Dynamics and Articulatory Phonology," in *Mind as Motion*, 175-194.

[12] Jeffrey Elman, "Language as a Dynamical System," in *Mind as Motion*, 195-226.

[13] M. T. Turvey and Claudia Carello, "Some Dynamical Themes in Perception and Action," in *Mind as Motion*, 373-402.

[14] Scot Kelso, *Dynamic Patterns: The Self-Organization of Brain and Behavior* (Cambridge: MIT Press, 1995).

[15] Hubert Dreyfus, "Merleau-Ponty and Recent Cognitive Science," in *The Cambridge Companion to Merleau-Ponty*, ed. Taylor Carman (Cambridge: Cambridge University Press, 2005), 129.

[16] Dreyfus, "Merleau-Ponty," 130.

infinite. That is why the robot, built according to the principles of Classic Cognition, does not solve the Frame Problem.[17]

In essence, EMT's critiques against the Representationalist Theory of Mind resort to: 1) the dense agent-environment coupling and 2) the absence of cognitive modules that communicate with each other by means of representations.[18] The brain, the body and the environment are equal partners in shaping the intelligent behavior. Consequently, how do we determine that only the neural factors are representational in nature? Secondly, in order to distinguish all the modules of the cognitive system, we have to a) identify the causal role of each cognitive module and b) explain the systemic properties in terms of system's parts features.[19] But the more intricate the relation between brain, body and environment is, the more difficult is to ascribe systemic properties to the system's parts.

Empirically, the neurobiologist Walter Freeman has proven that in case of perception, the aim of the cognitive system is not to create a faithful copy of stimuli. In his studies regarding the neurobiology of olfactory system of rabbits, he has remarked the phenomenon of variance of the neural patterns in the olfactory bulb (the so called AM – amplitude modulation patterns), given that the stimuli were the same. If the cognitivist thesis is true, the AM patterns should covary with the stimuli. The AM patterns variation depends on context, history and significance – whether the stimuli are associated either with reward or with punishment.[20] The AM patterns are not imposed from outside, they are created by the brain according to its own principles of self-organization.[21] When the rabbit perceives significant stimuli, Freeman has observed, strong bursts of energy cross

[17] Daniel Dennett, "Cognitive Wheels: The Frame Problem in Artificial Intelligence," in *The Robot's Dilemma: The Frame Problem in Artificial Intelligence*, ed. Zenon Pylyshyn (Norwood: Ablex, 1987), 41-65; Hubert Dreyfus, "Why Heideggerian AI Failed and How Fixing It Would Require Making It More Heideggerian," in *The Mechanical Mind in History*, eds. Philip Husbands, Owen Holland, and Michael Wheeler (Cambridge: MIT Press, 2008), 331-371.

[18] Michael Wheeler, "Two Threats to Representation," *Synthese* 129 (2001): 211-231.

[19] Wheeler, "Two Threats," 224.

[20] Walter Freeman, *How Brains make Up Their Minds* (New York: Columbia University Press, 2000), 77.

[21] "They cannot be representations of odorants because it is impossible to match them either with stimuli or with pulse patterns from receptors that convey stimuli to the cortex. It is also impossible to predict in detail the patterns that are constructed in the bulb from the patterns of receptor activation, because the constructions are by chaotic dynamics. They cannot be information because that is discarded in the spatial integration by divergent-convergent pathways. They are unique to the history of the individual, arising out of the past experience that shaped the synaptic connections in the bulbar neuropil." (Freeman, *How Brains*, 89-90.)

the nervous system. These states tend toward an energy minimum, which, in dynamic systems theory language, is called attractor. The system's entire activity can be seen as a transition from one attractor to another. The totality of the states tending toward the same attractor forms the attractor's basin. The brain develops basins of attraction for each significant class of inputs. Other experiences tend to integrate these basins of attraction, forming attractor landscapes. These landscapes govern the selection of the appropriate behavioral answer.[22]

The preemption argument

Some authors point out that Brooks and Co., by explaining gait development or limb movements, refer only to the online behavior which is generated by an ongoing interaction with the environment.[23] But there are offline actions, such as imagining counterfactual situations, planning vacations, arranging objects by their value etc., which, given the absence of direct environmental stimuli, require complex mental representations.[24] That is why Andy Clark argues[25] in favor of a hybrid version of EMT, which, on the one hand, will capture the dynamic aspects of agent-environment coupling using the concepts of dynamic theory, such as control parameters, collective variables, differential equations etc., and, on the other hand, will explain the offline actions by means of the classical cognitive sciences terms (representation, cognitive module, syntax etc.).

Fred Keijzer[26] argues that it is not necessary to postulate mental representations in order to explain 'representation-hungry' actions. He admits that there are internal states that determine the emergence of behavior, but the ontology of these states is essentially different from that of representations. There are two key concepts in Keijzer's dynamical explanation of behavior, namely, the control parameters picking out the conditions determining the evolution of the system in its space state and the order parameter describing the pattern of temporary order reached by the system. The control parameters, like, for example, the external stimuli, do not impose a certain order, but determine the system to construct a pattern as response according to its own principles of order.

---

[22] Freeman, *How Brains*, ch. 4.

[23] Wheeler, "Two Threats," 214.

[24] Andy Clark and Josefa Toribio, "Doing Without Representing?" *Synthese* 101, 3 (1994): 419-420.

[25] Clark, *Being There*, 126.

[26] Fred Keijzer, *Representation and Behavior* (Cambridge: MIT Press., 2001); Fred Keijzer, "Representation in Dynamical and Embodied Cognition," *Cognitve Systems Research* 3 (2002): 275-288.

Keijzer identifies the control parameters of the anticipatory behavior[27] using the comparison with the morphogenesis. Morphogenesis explains how, starting from an egg cell, an organism develops itself to full maturity. How does the cell know that after some divisions it must transform itself in a bony cell, epidermis etc.? The image of the cell as blueprint – complete developmental model – is, according to Keijzer, wrong.[28] The cell does not contain in itself the entire plan of organism's maturation, the extra-cellular factors being as important as the cell in shaping the organism. The cell is a specific internal control parameter (ICP); it guides the organism's evolution in its space-state, being at the same time part of the organism's dynamics. It pushes the organism to follow a certain 'epigenetic path.' Considering that the ICPs of anticipatory behavior are located at the neural level, Keijzer argues that mental representations do not fit the new picture because the representations function as complete models for behavior, but the neural ICPs are just modulators in a dynamic process of producing order.[29]

In the classic cognition, as Fodor often emphasizes, the concept of mental representation plays the role of offering a rational, logic account of behavior.[30] The order of behavior is encapsulated in the syntax of mental representations. But this is an assumed order and, consequently, a homuncular one.[31] The neural ICPs do not already contain in themselves the order, but they are part of a mechanism of producing order that comprises also extra-neural elements. According to Keijzer, the development of the anticipatory behavior emerges from the codetermination

[27] This is his term for representation-hungry actions.

[28] Keijzer, *Representation and Behavior*, 209-210.

[29] "In contrast to representations ICPs are intrinsically connected to a regulatory network of which they are a part. Also, ICPs do not consist of models of the external circumstances. Such an interpretation goes right against the grain of the idea of a regulatory trajectory. The macroscopic order is newly generated; it is not already encoded within the organism's ICPs." (Keijzer, *Representation and Behavior* , 241.)

[30] Jerry Fodor and Zenon Pylyshyn, "Connectionism and Cognitive Architecture: A Critical Analysis", in *Connectionism: Debates on Psychological Explanation.* vol. 2, eds. Cynthia Macdonald and Graham Macdonald (Oxford: Blackwell, 1995), 99, 112-113 (originally published in *Cognition* 28 (1988): 3-71).

[31] Distinguishing between intrinsic and derivative properties, Searle criticizes the claim that the brain is a digital computer. (Cf. John Searle, *The Rediscovery of the Mind* (Cambridge: MIT Press, 1992), 209-219). According to the computationalist theory of mind, the hardware level, or the brain, does operate with symbols in accordance with syntactic rules. If it were true that, for example, syntax is intrinsic to physical world, then everything instantiates a syntax, even the wall behind us would instantiate the program Word Star. (Searle, *The Rediscovery*, 208-209). Notions like 'symbol,' 'syntax,' 'program,' 'bits,' etc. always point to an interpreter.

of many scales of organization (sub-neural, neural and psychological scales).[32] Each scale is characterized by proper processes, situated in specific space-time dimension, conditioned by specific environmental stimuli (sub cellular, cellular and macroscopic stimuli),[33] being also in circular causation in which the neural factors modulate the body movements and these movements through a feed-back network influence the neural activity timing. To identify the neural activity with mental representations would mean to focus too much on one particular aspect of causation, neglecting the entire causal network in which they are embedded.[34] Being model for action, the classic mental representations are prior to the action. The agent acts as an intermediary between representation and environment. An ICP does not have significance outside the process of generating the behavior.

Adopting the same anti-representationalist stance, Anthony Chemero formulates the preemption argument.[35] He indicates an experiment[36] with subjects that receive sticks of different lengths (in increasing and then decreasing sequences); they should imagine that using those sticks they can move objects at a distance. Faced with the question whether they can move the object in front of them with those sticks, the subjects give answers that can be analyzed according to an order parameter and a control parameter. The situation is 'representation hungry' because the subjects must predict the result of actions which have not yet taken place.[37] Consequently, the dynamical explanation preempts the representational one: "If one has the complete dynamical story, what is left to be explained?"[38]

Representational states within dynamic systems

Chemero's thesis is that a fully dynamical, hence non-representational, story of our actions is empirically possible. Is this a recommendable strategy? I doubt that this is the case. He explains cases of strong decouplability by resorting to a special kind of oscillators that can keep track of absent stimuli.[39] But the elimination of mental representations is made only at the cost of transforming complex cognitive

[32] Keijzer, *Representation and Behavior*, 223. See especially, the figure 5.2.
[33] The neural process takes place in a few milliseconds and the psychological one in seconds.
[34] Keijzer, *Representation and Behavior*, 241.
[35] Anthony Chemero, *Radical Embodied Cognitive Science* (Cambridge: MIT Press, 2009).
[36] Iris van Rooij, Raoul M. Bongers and W. (Pim) F. G. Haselager, "A Non-Representational Approach to Imagined Action," *Cognitive Science* 26 (2002): 345-375.
[37] Chemero, *Radical Embodied Cognitive Science*, 40-42.
[38] Chemero, *Radical Embodied Cognitive Science*, 73.
[39] Chemero, *Radical Embodied Cognitive Science*, 54, 57-58.

agent-environment interactions into simple coupled oscillations.[40] In the case of offline actions, there is something in the agent that stands for the absent stimuli. A form of representation is necessary, though one which is different from the classical one. Instead of eliminating the concept of 'representation' (as radical dynamicists do) or of splitting cognitive explanation in two separate discourses (as the adepts of the hybrid version do[41]), I think that a dynamical concept of 'representation' is a better option. A dynamical conception of 'representation' may offer the advantages of avoiding the preemption argument, of unifying the cognitive explanation and, by its interscalar account, dynamic systems tools for building more complex artificial intelligent systems.

My argument is that, although it excludes the classic notion of 'representation,' Keijzer's scalar explanation of anticipatory behavior still encourages a representationalist interpretation. According to folk psychology, the psychological level states (thoughts, desires, intentions etc.) have two fundamental features: intentionality (they are about something) and causal relevance regarding behavior. In what follows I will try to describe the dynamic concept of 'representation' following these two features – intentionality and causal relevance.

The dynamic theory posits a circular causation in which the level S(n-1) acts as control parameter modulating the emergence of order at the level S(n). In turn, S(n) functions as order parameter constraining the activity of the components of S(n-1).[42] Also, the elements of S(n) act as control parameters for the level S(n+1). See figure 1.

---

[40] Chemero himself realizes that there is something more to representations that could not be picked up with oscillations, momentum etc.: "Noneffective tracking, though, is not sufficient for registration. In fact, noneffective tracking could be accomplished just by causal connection and momentum. (…) In registration, there is a further distancing and abstraction. It requires detachment in that the subject must 'let go' of the object, stop tracking it (even noneffectively). The difference here is like that between knowing your niece will come out from under the other side of the table, and knowing that you won't see her again until next Thanksgiving." (Chemero, *Radical Embodied Cognitive Science*, 57). The latter situation is surely not a case to be explained by oscillators and tracking.

[41] The above sketched preemption argument concerns particularly the hybrid version because only this account splits the cognitive explanation in two different vocabularies (corresponding to the two kinds of behaviors), while one of them can, in fact, cover both domains.

[42] Modulation means here that upward causation in which the components of a system push the system to follow a certain developmental path, without being full-fledged blueprints that determine each step of that path. Constraining means here the downward causation in which the system as a whole imposes some general conditions upon the components.

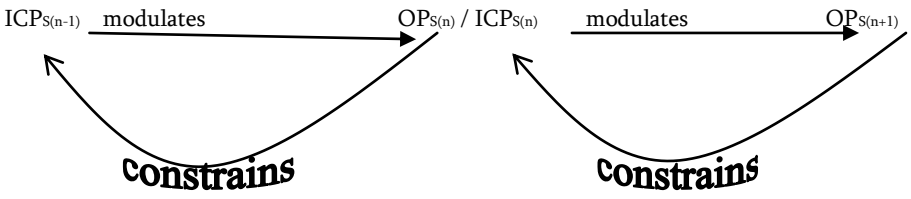

Fig. 1. Inter-scalar ratio between control parameters – ICPs – of the level S(n-1) and the order parameters – OPs – of the level S(n) and ICPs of the level S(n) and the OPs of the level S(n+1).

According to dynamic systems theory, the patterns can simultaneously play a double role: the role of downward causation (as order parameter) and upward causation (as control parameter).[43] For example, the cells make up the body, thereby setting its structure (there are different types of cells, for liver, heart, bones etc.), But the organism's existence as an autonomous entity guide their work.[44] The neuron (level S(n)) is, according to Kauffman,[45] the sequence of order emerging from the complex activity of hundreds of thousands of amino acids (level S(n-1)). At the same time, the neuron together with other many neurons act as ICPs modulating the emergence of anticipatory behavior at S(n+1). They do not dictate the order of behavior because it is not an already-given order; the order is generated during the brain-body-environment interaction.

In my view, the higher level order resulted from the complex brain-body-environment interaction is representational in nature. In dynamic system terms, the representational states are not the $ICP_{S(n)}$s but the $OP_{S(n+1)}$s.[46] The dynamic

[43] The dynamic system "is a set of quantitative variables changing continually, concurrently, and interdependently over quantitative time in accordance with dynamical laws described by some set of equations." (Robert Wilson and Frank Keil, eds., *The MIT Encyclopedia of the Cognitive Sciences* (Cambridge: MIT Press, 1999), 245.) The overall system activity emerges from the structural coupling of its parts, but the systemic properties are new in comparison to those of the parts; they represent more than their sum. Temporal rates of parts' activities are vital for the functioning of the whole and the system's overall state change, which occurs within the parameters set by the components, means the progress to a point in the space-state (that is, the space of all possible states of the system).

[44] Humberto Maturana and Francisco Varela, *The Tree of Knowledge. The Biological Roots of Human Understanding* (Boston: Shambhala, 1992), 87.

[45] Stuart Kauffman, *At Home in the Universe. The Search for Laws of Self-Organization and Complexity* (Oxford: Oxford University Press, 1995), 52.

[46] Mark Rowlands argues in his book *Body Language. Representation in Action* (Cambridge: MIT Press, 2006) that there is a sort of actions, called "deeds" that are intrinsically representational.

paradigm involves a significant change concerning the intentional nature of representational states: the basic forms of representations are not maps of reality implemented as such in the brain, but limit conditions or attractors constraining the cognitive system's evolution in its space state to reach its goals. On a certain threshold of complexity, the system develops stable attractors and attractor landscapes that could be interpreted as standing for something outside the system.

In the contemporary debate around representations, Haugeland's definition of this term is considered orthodoxy. Thus, for an organism to be credited as having representations:

> 1. It must coordinate its behaviors with environmental features that are not always 'reliably present to the system.'
> 2. It copes with such cases by having something else 'stand in' for those features and guide behavior.
> 3. The 'something else' is part of a more general representational scheme that allows the standing in to occur systematically and allows for a variety of related states.[47]

The dynamic concept of 'representation' meets the first requirement simply by the fact that many organisms react to stimuli that are absent. There are cases of dense agent-environment coupling, where we could not find enduring internal states guiding these couplings. For example, in the case of Watt governor,[48] the dynamic codependence between the speed of the engine and the steam pressure can be picked up in non-representational terms. The gill withdrawal reflex in the sea slug *Aplysia* can be studied as a chemical process in which the presynaptic motor neuron release less neurotransmitter due to a blockage of the calcium channels.[49] In such systems there are no enduring states.[50] However, there are more complex cognitive systems whose adaptive behavior is based on past experiences, that is, on enduring states.[51] The current states of the system are

[47] John Haugeland, "Representational genera," in *Philosophy and Connectionist Theory*, eds. William Ramsey, Stephen P. Stich, and David E. Rumelhart (Hillsdale: Erlbaum, 1991), cited in Clark *Being There*, 144.

[48] Thimothy van Gelder, "Dynamics and Cognition," in *Mind Design II. Philosophy, Psychology, Artificial Intelligence*, Revised and enlarged edition, ed. John Haugeland (Cambridge: MIT Press, 1997), 422-429.

[49] Arthur Markman and Eric Dietrich, "In Defense of Representation," *Cognitive Psychology* 40 (2000): 148.

[50] Markman and Dietrich, "In Defense of Representation," 148.

[51] "(…) systems that learn and make use of prior behavior have some enduring states that allow the system to react to new situations on the basis of past experience." (Markman and Dietrich, "In Defense of Representation," 148.)

directly coupled to the stimuli, but to something that is not reliable present (the past experiences).

The second requirement states something more than simply reacting to absent stimuli. It asserts the ontological condition that the representational cognitive systems should have specific internal states playing a representational role. My thesis is that the representational states are not explicit, full-fledged representations of absent stimuli, but an end-state, a goal state that guide the behavior of the more evolved systems. For example, the desire to drink a cup of milk triggers a certain action. This desire is a representational state, because it is an internal event that stands for an absent object. The desire to drink milk is an internal event, but "However, as soon as the action is initiated and I am on my way to the fridge for that glass of milk, my action is an ongoing affair that involves the ongoing scanning of my visual environment, using the results to adjust my movements and so on."[52] In this case, the desire is the permanent internal condition modulating the agent's interaction with the world toward a certain goal. In this case we can observe that there are internal states that do not emerge from the online coupling; they are prior to the coupling and more than that, they guide this coupling.

Hubert Dreyfus believes that such actions are purposive "without the agent entertaining a goal."[53] By means of the concept of 'optimal grip' (Merleau-Ponty) he explains purposive actions as a process of searching for equilibrium states. We are autonomous beings in need to cope with the environment. As such, "we are constantly 'motivated' to move so as to achieve the best possible grip on the world."[54] There is no need of explicitly representing goals because, when the deviation from optimal interaction occurs, the agent tends to look for a better grip on the environment, without knowing what she is looking for.

I think that Dreyfus' account neglects the fact that there could be two kinds of goals. His thesis is valid in situations in which, for example, I sit in an uncomfortable chair and keep moving almost unconsciously till I get the best position. In this case, there is no explicit goal inside the agent. But if I want a glass of milk, it could be either because I am hungry (and I happen to have only milk in

[52] Fred Keijzer and Maurice Schouten, "Embodied Cognition and Mental Causation: Setting Empirical Bounds on Metaphysics," *Synthese* 158 (2007): 119. For Keijzer and Schouten this desire is not a representational state, but simply a "psychological entity." (Keijzer and Schouten, "Embodied Cognition," 118).

[53] Dreyfus, "Merleau-Ponty," 138.

[54] Dreyfus, "Merleau-Ponty," 137.

the fridge) or because I prefer milk. In the latter case, I have an explicit goal, a mental state with specific semantic content: it stands for an object.

Regarding the third requirement, on my approach there is no gap between online and offline actions. For less evolved cognitive agents, the goal (or, in dynamical terms, attractor) means a release from tension. But as the system learns to use past experiences, in order to lower the tension, it develops more complex enduring states. According to Markman and Dietrich,[55] this is the second level of representation (the first one is that of simple mediating states). In the case of complex human cognition, the attractor shapes not just the tendency toward a stable, ordered energy state, but also specifies the semantic content of that end state. When it comes to semantic content, the state is systematically connected to other states, by means of this content. The evolved cognitive systems develop stable attractor landscapes that can be decomposed, rearranged, modeled etc.[56]

Developing further Keijzer's interscalar account of anticipatory behavior, we may find ways of improving the performances of the artificial intelligent systems. We observe that, on the interscalar account, the causal role of the representational states manifests itself in two ways: first, by influencing the neural activity (as order parameter) and secondly, by modulating the metarepresentational level (as control parameters). For example, Herbert, the robot built by Brooks in the 90s, could collect cans by running in parallel simpler actions, such as walking, avoiding obstacles, identifying cans. Its behavior is modulated by an electronic internal parameter. At first sight, the action is non-representational. However, if Herbert has to pick up only the valuable cans,[57] it should develop a new level of action which, in my view, is meta-representational. Herbert has to collect items which have something in common, but not at the level of physical properties. Therefore, Herbert operates not with a purely physical input, but with certain representations (of valuable cans). Only the meta-representational level enables us to grasp the presence of representational states. Herbert's new behavior is presented in figure 2 as follows:

[55] Markman and Dietrich, "In Defense of Representation," 148.

[56] See also, for example, Andy Clark, *Supersizing the Mind. Embodiment, Action, and Cognitive Extension* (Oxford: Oxford University Press, 2008), 27.

[57] This is, according to Clark and Toribio ("Doing Without Representing?"), a representation hungry action.

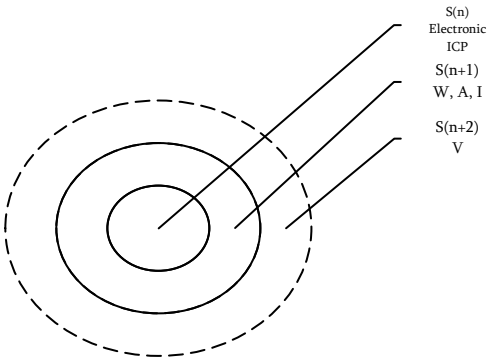

Fig. 2. The scalar structure of Herbert's offline behavior. W = walking, A = avoiding, I = identification, V = identification of valuable cans

The patterns W, A and I represent the order parameters resulting from robot-environment coupling. But they function as control parameters for the pattern V. At the same time, without the level S(n+2), it is not possible to grasp the relational (hence, intentional) nature of W, A and I. The lack of metarepresentations is perhaps the reason why Brooks has defended a nonrepresentational theory of behavior. The metarepresentational level is not condition of possibility for the simpler actions W, A and I as such, but it enables the agent to 'see' its relation to the world. The action V is not possible unless the robot operates with its own perspective on things.

We would not see our relation to the world, if we had not a meta-representational level. As Wittgenstein would say, a representation points to a state of affairs, but not to its own relation with that state. The relation shows itself; it is not part of the representation. Actually, we see the intentional relation by means of metarepresentations, that is, at the level where the subject sees itself as being in relation with the world. Here, we come across the essential meta-representation, the self.

By my argument I try to answer the question whether we can ascribe representational states to the dynamic explanations of mind. However, according to Ramsey,[58] this is a trivial achievement because, as Dennett has showed, [59] even a stone could be described through the intentional stance. In Ramsey's view, the non-trivial questions regarding representations refer to whether there is any explanatory benefit in describing the cognitive processes in representational terms

[58] William Ramsey, *Representation Reconsidered* (Cambridge: Cambridge University Press, 2007), 33-4.

[59] Daniel Dennett, *Brainstorms* (Cambridge: MIT Press, 1978).

and whether there are internal states that function as representations in a robust and recognizable manner.[60] In reply: yes, we have meant to argue that a representational interpretation of the dynamic explanation of mind is possible; but this is not a trivial achievement, because it offers the advantage of unifying the field of the cognitive explanation. If the explanation were wholly dynamical, it would exclude the offline behaviors, and if it were a hybrid version, it could not explain how the two discourses (dynamic and classic computational) do stay together.

In the dynamic paradigm I propose here, the representational states suffer significant changes concerning their ontological status. In the classic cognitive paradigm mental states are symbols implemented as such in the brain. Hence, their intentionality, their aboutness, is external to them. In the new dynamic paradigm the intentionality of mental states derives from the self-organizing processes of an embodied and embedded agent. The mind spontaneously tends towards its attractors; these attractors are not copies of the stimuli, but express the internal energy equilibrium states to which the system is driven by its own principles of organization, given the impact of the environmental stimuli. Does it follow from this that the intentional objects are just constructions of the cognitive system? In the new paradigm the relation between the intentional state and its object could not be a linear one, namely, from stimulus to representation. The system generates dynamic trajectories in its space-state and stabilizes them in attractor landscapes as responses to the environmental perturbations; in turn, these responses modulate the perceptual activity of the system and so on. Obviously, the difference between classic and dynamic explanation reflects the difference between the textual paradigm and the processual one.[61] Mental states are not static symbols encoding semantic content, but topological entities, evolutions in space-time that constitute their object. From another perspective, the same difference could be understood as the difference between the heteronomic approach where the mental representations are faithful copies imposed as such by the stimuli and the approach based on self-organization, where the stimuli are just affordances that help the organism to enact its own sensory-motor domain of significance.[62]

---

[60] Ramsey, *Representation Reconsidered*, 34.

[61] Clark has made this remark in his book *Associative Engines. Connectionism, Concepts and Representational Change* (Cambridge: MIT Press, 1993), 8, in the context where he discusses the virtues of the connectionist networks.

[62] Cf. James Gibson, *The Ecological Approach to Visual Perception* (Boston: Houghton- Mifflin, 1979); Varela, Thompson, and Rosch, *The Embodied Mind*; Evan Thompson, *Mind in Life* (Cambridge: Harvard University Press, 2007); Dreyfus, "Why Heideggerian AI Failed."

DEBATE

PROBLEMS WITH THE DISPOSITIONAL TRACKING THEORY OF KNOWLEDGE

Ben BRONNER

ABSTRACT: Rachel Briggs and Daniel Nolan attempt to improve on Nozick's tracking theory of knowledge by providing a modified, dispositional tracking theory. The dispositional theory, however, faces more problems than those previously noted by John Turri. First, it is not simply that satisfaction of the theory's conditions is unnecessary for knowledge – it is insufficient as well. Second, in one important respect, the dispositional theory is a step backwards relative to the original tracking theory: the original but not the dispositional theory can avoid Gettier-style counterexamples. Future attempts to improve the tracking theory would be wise to bear these problems in mind.

KEYWORDS: defining knowledge, tracking theory, safety and sensitivity, Robert Nozick, Rachel Briggs, Daniel Nolan

Rachel Briggs and Daniel Nolan amend Nozick's tracking theory of knowledge.[1] Nozick claims that *S* knows that *p* if and only if:

(1) *p* is true.

(2) *S* believes that *p*.

(3) If *p* had not been true, then *S* would not have believed that *p*.

(4) If *p* had been true, then *S* would have believed that *p*.[2]

Briggs and Nolan suggest replacing (3) and (4) with:

(3*) *S* is disposed to not believe that *p* in the circumstance where *p* does not obtain.

(4*) *S* is disposed to believe that *p* in the circumstance where *p* obtains.

As they demonstrate, Briggs and Nolan's modified, dispositional tracking theory can deal with several counterexamples to the original tracking theory.

The dispositional tracking theory faces significant problems of its own, however. John Turri notes that satisfying the conditions of the theory is not necessary for knowledge.[3] He supplies several examples where a subject has knowledge and yet does not meet condition (3*). Despite this problem, Turri

[1] Rachel Briggs and Daniel Nolan, "Mad, Bad and Dangerous to Know," *Analysis* 72, 2 (2012): 314-316.

[2] Robert Nozick, *Philosophical Explanations* (Cambridge: Harvard University Press, 1981).

[3] John Turri, "Stumbling in Nozick's Tracks," *Logos & Episteme* 3, 2 (2012): 291-293.

remarks that Briggs and Nolan make "modest progress along the path Nozick helped to pioneer." Certainly others will attempt to progress further along that path, and the examples Turri supplies will help in this project. Further progress may be limited, however, if nothing more is said. First, it is not simply that satisfaction of the conditions of the dispositional tracking theory is unnecessary for knowledge – it is insufficient as well. Second, in one important respect, the dispositional theory is a step *backwards* relative to the original tracking theory: the original but not the dispositional theory can avoid Gettier-style counterexamples.

To illustrate these problems, first consider one of the cases Briggs and Nolan describe.

> DANGEROUS TO KNOW Adolf believes, having studied his symptoms and consulting with medical experts, that he has a rare brain condition that is fatal before the age of 5 years in 99.99% of cases. In fact, he is right; he's one of the one in 10,000 who have the disease but survive. However, in most nearby possible worlds here he has the disease, he dies before the age of 5 years – long before he entertains the proposition that he has the disease. Therefore, it is not true that if Adolf had had the disease, he would have believed he did. (He might well have died at 5 years.) So Adolf's belief is not safe. Nonetheless, Adolf knows he has the disease.

Though Adolf does not satisfy the conditions of the original tracking theory, he does satisfy those of the dispositional theory. After all, "Adolf has the disposition to take the word of reliable and informed authorities about the state of his health."

Now consider the following alteration of Adolf's case. Everything is as before, except: The 'medical experts' are actors and Adolf is unwittingly the subject of an elaborate hoax funded by his wealthy arch nemesis, Olaf. Not only has Olaf hired these actors, but he has tricked Adolf into thinking that the latter is manifesting symptoms of the disease in question. Unknown to Adolf, what he takes to be symptoms of the disease are in fact symptoms of the various drugs that Olaf has been slipping into Adolf's food. Unknown to both Adolf and Olaf, the former truly has the disease, but is asymptomatic.

This seems to be a Gettier-style case where Adolf has a lucky true belief but no knowledge. Yet all four of Briggs and Nolan's conditions seem to be satisfied. In particular, all of Adolf's dispositions are the same as in the original case. From Adolf's point of view, the situation is indistinguishable from that described by Briggs and Nolan. There is no difference between Briggs and Nolan's example and the modified example that could have psychologically impacted upon Adolf, changing his dispositions. It follows that satisfaction of the conditions advanced by

Briggs and Nolan is not sufficient for knowledge. It is worth emphasizing that Nozick's tracking theory produces the correct verdict, as condition (3) is not satisfied: if Adolf had not had the disease, he still would have believed that he did. In this regard, moving from the original to the dispositional tracking theory is a step in the wrong direction – it is not simply that the dispositional theory makes "modest progress" before encountering the problem noted by Turri. Future attempts to improve the tracking theory would be wise to bear these additional problems in mind.[4]

[4] Special thanks to Dorit Ganson for helpful comments.

LOTTERIES, PROBABILITIES, AND PERMISSIONS

Clayton LITTLEJOHN

ABSTRACT: Thomas Kroedel argues that we can solve a version of the lottery paradox if we identify justified beliefs with permissible beliefs. Since permissions do not agglomerate, we might grant that someone could justifiably believe any ticket in a large and fair lottery is a loser without being permitted to believe that all the tickets will lose. I shall argue that Kroedel's solution fails. While permissions do not agglomerate, we would have too many permissions if we characterized justified belief as sufficiently probable belief. If we reject the idea that justified beliefs can be characterized as sufficiently probably beliefs, Kroedel's solution is otiose because the paradox can be dissolved at the outset.

KEYWORDS: epistemic obligation, evidence, justification, Thomas Kroedel, lottery paradox, probability

Thomas Kroedel[1] argues that if we assume that we are permitted to believe *p* iff we are justified in believing *p*, we can easily solve a version of Kyburg's[2] lottery paradox.[3] Here is the set up. You know that there is a large and fair lottery. Only one ticket can win and the odds of any ticket winning is the same as the odds of any other ticket winning. It seems to him that:

(1-J) For each ticket, you are justified in believing that it will lose.

The paradoxical conclusion that is supposed to follow from (1-J) is that:

(2-J) You are justified in believing that all the tickets will lose.

It seems that (2-J) is false. What to do?

Kroedel suggests that (1-J) and (2-J) are equivalent to:

(1-P) For each ticket, you are permitted to believe that it will lose.

---

[1] Thomas Kroedel, "The Lottery Paradox, Epistemic Justification, and Permissibility," *Analysis*, 72, 1 (2012): 57-60.

[2] Henry Kyburg, *Probability and the Logic of Rational Belief* (Middletown: Wesleyan University Press, 1961).

[3] For arguments in support of the view that justified beliefs are permissibly held beliefs (but not beliefs you're obligated to have), see Clayton Littlejohn, Justification and the Truth-Connection (Cambridge: Cambridge University Press, 2012) and Mark Nelson, "We Have No Positive Epistemic Duties," *Mind* 119 (473): 83-102.

(2-P) You are permitted to believe that all the tickets will lose.

He says that (1-P) is ambiguous between a narrow- and wide-scope reading. On one reading, (1-P) is false. On one reading (1-P) is true, but it does not entail (2-P).

Let *p*, *q*, *r*, etc. be propositions about particular tickets losing. Let PB*p* be the sentence "It is permissible for you to believe *p*." Here is the first reading of (1-P), a narrow-scope reading:

(N1-P) PB*p* & PB*q* & PB*r*

Kroedel says that (N1-P) is true. He is right that (N1-P) does not entail (2-P) because permissions do not agglomerate. If permissions agglomerate, whenever you are permitted to ϕ and permitted to ψ, you would be permitted to ϕ and ψ. If, say, we were sharing a cake and you were permitted to take one half or take the other half, you would thereby be permitted to take the whole thing if permissions agglomerated. Since you can be permitted to take part without thereby being permitted to take it all, permissions do not agglomerate. Having two permissions and using one can thereby lead you to lose the other. Likewise, being permitted to believe *p* and being permitted to believe *q* does not mean that you are permitted to believe both *p* and *q* (or believe the conjunctive proposition that *p* and *q*).

Contrast (N1-P) reading with this wide-scope reading of (1-P):

(W1-P) P[B*p* & B*q* & B*r*]

While (W1-P) does entail (2-P), Kroedel argues that (W1-P) is false. He says that the problem with (W1-P) is that it is plausible that *if* you are permitted to hold several beliefs then you are thereby permitted to have a single belief that is the conjunction of those contents. It would be if this closure principle were true:

(CP) If P[B*p* & B*q* & B*r*], then PB[*p* & *q* & *r*]

It is, as he notes, highly implausible that you are permitted to believe the conjunctive proposition [*p* & (*q* & *r*)], so if the permission to believe both *p* and *q* (together) carries with it the permission to believe the conjunction (*p* & *q*), we have some reason to think that (W1-P) is false.

Kroedel thinks that to solve the lottery paradox, we should say that beliefs concerning lottery propositions can be justified but there is a limit as to how many such beliefs we can permissibly form. He thinks that there is a reading of (1-P) that is true, (N1-P) and urges us to reject (2-P). Why should we accept (N1-P) or (1-J)? His suggestion is that the high probability of a proposition is sufficient for the permissibility of believing it (forthcoming: 3). What is wrong with (2-P)? The problem cannot be that if you were permitted to believe that all the tickets will lose you will thereby believe something you know is false. The lottery might not

have a guaranteed winner and (2-P) still seems false. Perhaps the reason that (2-P) is false is just that the probability that all the tickets will lose is too low. It might seem that Kroedel's solution should work given the following principle linking justification to probability:

(PJ) PB*p* iff the probability of *p* on your evidence is sufficiently high.

While (PJ) would (if true) explain why (1-P) is true and (2-P) is not, I think Kroedel has to reject (PJ). If he does that, his solution becomes otiose. We can dissolve the paradox by denying (1-J) and (N1-P) rather than worrying about whether these false claims entail further false claims.

To see this, notice that since Kroedel is committed to (N1-P) and denying (2-P), he is committed to the following claim:

(*) For a lottery with n tickets (assuming that n is suitably large number), you are not permitted to form n beliefs that represent tickets in the lottery as losers but for some number m such that $n > m > 0$, you are permitted to form m beliefs that represent tickets in the lottery as losers.

Suppose you had a ticket for a lottery with 1,000,000 tickets. You know that this ticket, ticket #1, is not terribly likely to win. Let us suppose that in a lottery this size, you are permitted to believe that at least one ticket will lose. You decide to make use of this permission and believe ticket #1 will lose. You know (or should know!) that forming this belief does not change the probability that ticket #2 will lose. Since you were permitted to believe it before believing ticket #1 will lose and its probability remains unchanged when you add that first belief to your belief set, you can add the belief that ticket #2 will lose without being compelled to abandon your belief that ticket #1 will lose. (It is not as if adding the belief that ticket #2 will lose to your belief set forces you to lower the probability that ticket #1 will lose.) Given (PJ), it seems you can permissibly believe both that ticket #1 and ticket #2 will lose. We can apply the same reasoning again and you can permissibly add the belief that ticket #3 will lose without having to abandon previously formed beliefs about losing tickets. Repeat. At some point, (PJ) says that you can add some number of beliefs about losing lottery tickets greater than m without impermissibly adding beliefs to your stock of beliefs. At this point, (*) says that you formed more beliefs than you are permitted to form. You cannot consistently endorse (PJ) and (*).

If Kroedel denies (*), he has to either accept (2-P) or reject (N1-P). If he does that, he has to abandon his proposed solution. If instead he retains (*) and denies (PJ), he has to deny one of the following:

(High) PB*p* if the probability of *p* on your evidence is sufficiently high.

(Low) ~PB*p* if the probability of *p* on your evidence is sufficiently low.

If he denies (Low), it is not clear why we should reject (2-P). If he denies (High), it is not clear what motivation there is for (N1-P). If forced to choose, it seems rather obvious that he should deny (High) rather than (Low). If he rejects (High), he can dispense with the lottery paradox much more quickly. If it is possible for the probability of *p* to be sufficiently high on your evidence and for you to be obligated to refrain from believing *p*, I cannot see what would be wrong with saying that your obligation is to refrain from believing lottery propositions. The only thing they have going for them from the epistemic point of view is their high probability.

There is a common objection to (High) and to (1-P). Many people think that you cannot know that a ticket in a lottery with 1,000,000 tickets will lose. Suppose this is so and suppose that you know that you cannot know that the ticket you have is a loser. If you believe that your ticket is a loser and know that you cannot know that your ticket is a loser, this is how you see things:

(1) This ticket will lose, but I do not know that it will.

If you believed such a thing, you would be deeply irrational. While you cannot justifiably believe (1), the probability of (1) on your evidence, however, is quite high. So, (High) says that it is permissible to believe (1). So, (High) is mistaken. If (High) is rejected, this should take some of the sting out of denying (1-J) and (N1-P).

One reason to think that you cannot justifiably believe lottery propositions is that you cannot justifiably believe what you know you cannot know.[4] Lottery propositions are known unknowns. There is a further reason to think that it would be better to solve the lottery paradox by denying (1-J) and (N1-P) than to accept these and reject (2-J). It seems you cannot have proper warrant to assert:

(2) Your ticket will lose.

You might explain why it would be improper to assert (2) given only knowledge of the odds that the ticket will lose on the grounds that you cannot know (2) and cannot have warrant to assert what you do not know.[5] For various reasons, some object to this suggestion on the grounds that there are propositions that the speaker does not know that the speaker does have sufficient warrant to

[4] Alexander Bird, "Justified Judging," *Philosophy and Phenomenological Research* 74 (2007): 81-110; Jonathan Sutton, "Stick to What You Know," *Nous* 39 (2005): 359-96.

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

assert. (Maybe you do not know that the building you see on the hillside is a barn, but if you do not know this just because you are in fake barn country, it is not obvious that you do not have sufficient warrant to assert that the building is a barn.) Suppose that knowledge is not necessary for warranted assertion. If knowledge is not necessary for warranted assertion, what is? Douven[6] argues that reasonable belief is necessary and sufficient for warranted assertion. Given Douven's account, we have to reject (N1-P) and (1-J) to explain why you do not have sufficient warrant to assert (2). You cannot properly assert (2) because you cannot justifiably believe (2). Thus, (1-J) is mistaken. Paradox dissolved.

Someone might say that there is an alternative approach available to Kroedel, one on which we modify (High) as follows:

> (High*) PBp if the probability of *p* on your evidence and what you permissibly believe is sufficiently high.

Armed with this, he might argue that as you add more and more lottery beliefs to your set of beliefs, there will come a point at which the probability that some ticket is a loser will be too low for you to add another belief to your belief set. The thought seems to be that it would be impermissible to add a belief to your belief set because the probability that the remaining tickets will lose *on your beliefs* falls further and further the more beliefs you form. Eventually, it becomes impermissible to add more.

For this move to work, we would have to say that whether one has sufficient propositional justification to believe a lottery proposition depends upon the probability of that proposition conditional on what you justifiably believe and on what your evidence is. The oddity of this response, however, seems to be that the minimally rational agent knows that the probability of each ticket turning out to be a loser is the same and remains invariant however many lottery beliefs the subject forms. Yet, (High*) suggests that the reason that you don't have the permission to believe some lottery proposition is that the probability of some ticket turning out to be a loser has dropped below some threshold.

Someone might instead suggest that the problem has to do with the agglomeration of risk. With each belief comes additional risk and with each additional risk taken, you get closer and closer to a level of risk taking that is unacceptable.

The problem with this reformulation of the solution is that it succumbs to a version of the wrong kind of reasons problem. We have already established that

---

[6] Igor Douven, "Assertion, Knowledge, and Rational Credibility," *Philosophical Review* 115 (2006): 449-85.

the additional lottery beliefs you form are no riskier than other beliefs you have formed. For Kroedel's solution to work, we have to assume that some beliefs in the set of lottery beliefs are permissibly held. While adding more and more beliefs means that the probability of the set of lottery beliefs will include a falsehood increases, the fact that adding a belief to a set increases the probability that the set contains a falsehood is not the sort of reason that counts against adding that particular belief to the set. Adding the preface-like belief to your belief set (e.g., that something you believe is mistaken) guarantees that the set of beliefs you have contains a falsehood, but *that* fact does not constitute a decisive reason to remain agnostic about your own fallibility or to believe that you are not in error in any of your beliefs.

If (High) should be abandoned, (High*) is not a suitable alternative. If neither (High) nor (High*) is acceptable, it is difficult to see how a solution along the lines proposed by Kroedel could work. If no modification of (High) or (High*) is acceptable, we can reject (High) and (High*). Having done that, there is no problem left for us to solve.

NOTES ON THE CONTRIBUTORS

Robert Albin is Senior Lecturer in Philosophy and Ethics at The Department of Liberal Arts at Sapir College, Israel. His Research interests are in the fields of philosophy of culture and applied ethics, with a special interest in media ethics. He is the author of *A Chronicle of the Decline of Rationality: Ethics in the Practice of Journalism* (2004, in Hebrew). His publications include: "Journalists as Agents of Cultural Change: From Rationality Back to Nature" (*International Journal of Applied Philosophy*, 2007), "Modern Technology as a Denaturalizing Force" (*Poiesis and Praxis*, 2007), and "Moral Dilemmas in Business Ethics: From Decision Procedures to Edifying Perspectives" (*Journal of Business Ethics*, 2006, with Yotam Lurie). Contact: albin@sapir.ac.il.

George Bodi, PhD, is scientific researcher for the Romanian Academy – Iaşi Branch, Institute of Archaeology. He is currently beneficiary of a post–doctoral scholarship within *The Knowledge Based Society* project developed by the Romanian Academy, with the individual research project *The Constitution of Knowledge in the Archaeology of Romanian Prehistory*. His main research interests are situated within the field of landscape archaeology and archaeological theory. Contact:georgebodi@gmail.com.

Bogdan Boghițoi is a postdoctoral researcher at the Faculty of Philosophy of the University of Bucharest. His interests lie chiefly in the epistemological problems raised by the various theories of psychology and cognitive science. He is the author of *Mintea familiară* (*The Familiar Mind*) (Editura Trei, 2010), which aims at clarifying the structure and scientific credentials of folk psychology. Contact: bogdan.boghitoi@gmail.com.

Anthony Booth is lecturer in philosophy at the University of Sussex, United Kingdom. His research interests are in the philosophy of mind, ethics, and epistemology more generally, and on the ethics of belief more specifically. Contact: A.R.Booth@sussex.ac.uk.

Ben Bronner recently received his BA in philosophy and neuroscience from Oberlin College. His main interests are in the philosophy of mind and normative and applied ethics. He is the co-author of "Visual Prominence and Representationalism" (with Todd Ganson, *Philosophical Studies*, forthcoming) and "Burge's Defense of Perceptual Content" (with Todd Ganson and Alex Kerr,

Philosophy and Phenomenological Research, forthcoming). Contact: benj.bronner @gmail.com.

Jasper Doomen is a lecturer in law at Leiden University and has previously been active in the same capacity at Utrecht University. He holds an M.A. in Philosophy (Leiden University, 2003) and a J.D. (Utrecht University, 2005). His publications are mainly in the field of philosophy and law. Contact: jasperdoomen@yahoo. com.

Dale Jacquette is ordentlicher Professor für Philosophie, Abteilung Logik und theoretische Philosophie (Senior Professorial Chair in Logic and Theoretical Philosophy), at Universität Bern, Switzerland. He is the author of numerous articles on logic, metaphysics, philosophy of mind, and aesthetics, and has recently published *Philosophy of Mind: The Metaphysics of Consciousness* (Continuum, 2009), *Ontology* (McGill-Queen's University Press, 2002), *David Hume's Critique of Infinity* (Brill, 2000), and *Wittgenstein's Thought in Transition* (Purdue University Press, 1998). He has edited the *Blackwell Companion to Philosophical Logic* (2005), *Cambridge Companion to Brentano* (2004), and for North-Holland (Elsevier) the volume on *Philosophy of Logic* (2007) in the *Handbook of the Philosophy of Science* series. Contact: dale.jacquette@philo.unibe.ch.

Jimmy Alfonso Licon recently took his MA in Philosophy from San Francisco State University. He plans to apply to philosophy doctoral programs in the Fall. His philosophical interests are various issues in metaphysics and epistemology, and applied ethics. He published in *Logos & Episteme* "No Suicide for Presentists: A Response to Hales" and "Still No Suicide for Presentists: Why Hales' Response Fails." Contact: jimmylicon01@gmail.com.

Clayton Littlejohn is Lecturer with the Department of Philosophy at King's College London. His main areas of research are in epistemology and ethical theory. He is the author of *Justification and the Truth-Connection* (Cambridge University Press, 2012), a monograph on the internalism-externalism debate in epistemology. Currently, he is co-editing *Epistemic Norms* (Oxford University, forthcoming) with John Turri and writing *This is Epistemology* (Wiley-Blackwell, forthcoming), an introduction to epistemology. Contact: clayton.littlejohn@kcl. ac.uk.

Nicolás Lo Guercio is a graduate student in the University of Buenos Aires and a scholar in the National Center for Research in Science and Technology, Argentina. He is adjunt in Logic in the UNQui (National University of Quilmes), Argentina. His main areas of research include epistemology (peer disagreement and

epistemology of testimony) and philosophy of language (truth-relativism, the semantics/pragmtaics boundary). He is part of the research project "Contemporary debates on epistemology of testimony" (UBACyT, ANCyT and SADAF). Contact: nicolasloguercio@gmail.com.

Rik Peels is a post-doctoral researcher at the VU University Amsterdam, the Netherlands. His research interests are in ethics, epistemology, philosophy of religion, and philosophical theology. His most recent publications are "Tracing Culpable Ignorance" (*Logos & Episteme*, 2011), "Sin and Human Cognition of God" (*Scottish Journal of Theology*, 2011), and "Ignorance Is Lack of True Belief: A Rejoinder to Le Morvan" (*Philosophia*, 2011). Contact: mail@rikpeels.nl.

Noah Roderick is an Assistant Professor of English at Lourdes University in Sylvania, Ohio, United States. His teaching and research areas include the rhetoric of science, social epistemology, and minority language and literacy studies. His most recent publications include "Analogize This! The Politics of Scale and the Problems of Substance in Complexity-Based Composition," published in *Composition Forum* (Spring 2012) and "Beyond the Common Denominator: Exposing Semiotic (Dis)Unity in Mathematics Textbooks" (co-written with Travis Olson), which appeared in *The New Politics of the Textbook: A Project of Critical Examination and Resistance* (Sense Publishers: 2012). Contact: nroderic@ lourdes.edu.

Cristinel Ungureanu, PhD in philosophy at "Alexandru I. Cuza" University of Iaşi in 2010, is an independent researcher. His main philosophical interests are in the philosophy of mind and philosophy of psychology. He publishes on these issues in various philosophical journals and collective volumes. He has been awarded several research scholarships and fellowships from Centrul National de Burse, DAAD, and "Alexandru I. Cuza" University of Iaşi. Contact: c_ungureanu2003@ yahoo.com.

LOGOS & EPISTEME: AIMS & SCOPE

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

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

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

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

NOTES TO CONTRIBUTORS

1. Accepted Papers

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

2. Submission Address

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

3. Paper Size

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

4. Manuscript Format

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

5. Official Languages

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

6. Abstract

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

7. Author's CV

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

8. Review Process

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

9. Acceptance of the Papers

The editorial committee has the final decision on the acceptance of the papers. Papers accepted will be published, as far as possible, in the order in which they are

received and they will appear in the journal in the alphabetical order of their authors.

10. Responsibilities

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

11. Checking Proofs

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

12. Reviews

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

13. Property & Royalties

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

14. Permissions

Authors have the right to use their papers in whole and in part for non-commercial purposes. They do not need to ask permission to re-publish their

papers but they are kindly asked to inform the Editorial Board of their intention and to provide acknowledgement of the original publication in *Logos & Episteme*, including the title of the article, the journal name, volume, issue number, page number and year of publication. All articles are free for anybody to read and download. They can also be distributed, copied and transmitted on the web, but only for non-commercial purposes, and provided that the journal copyright is acknowledged.

Bun de tipar: 2012 • Apărut: 2012 • Format 17 × 24 cm

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

Volum elaborat în cadrul Proiectului POSDRU/89/1.5/S/56815.
Se distribuie GRATUIT.