THE JOURNAL OF PHILOSOPHY

VOLUME CVII, NUMBER 6
JUNE 2010

page

***Published by* The Journal of Philosophy, Inc.**

THE JOURNAL OF PHILOSOPHY

FOUNDED BY FREDERICK J. E. WOODBRIDGE AND WENDELL T. BUSH

Purpose: To publish philosophical articles of current interest and encourage the interchange of ideas, especially the exploration of the borderline between philosophy and other disciplines.

All communications to the Editors and Trustees and all manuscripts may be sent to Alyssa Timin, Managing Editor, at submissions@journalofphilosophy.org.

You may also visit our website at: www.journalofphilosophy.org

Published monthly as of January 1977; typeset by Dartmouth Journal Services, Waterbury, VT, and printed by The Sheridan Press, Hanover, PA.

All communication about subscriptions and advertisements may be sent to Pamela Ward, Business Manager, Mail Code 4972, 1150 Amsterdam Avenue, Columbia University, New York, NY 10027. (212) 866-1742

The Journal of Philosophy (ISSN 0022-362X) is published monthly by The Journal of Philosophy, Inc., MC4972, Columbia University, 1150 Amsterdam Avenue, New York, NY 10027. Periodicals postage paid at New York, NY, and other mailing offices.

POSTMASTER: Please send address changes to *The Journal of Philosophy* at MC4972, Columbia University, 1150 Amsterdam Avenue, New York, NY 10027.

ISSN 0022–362X

THE JOURNAL OF PHILOSOPHY
VOLUME CVII, NO. 6, JUNE 2010

MEASUREMENT-THEORETIC REPRESENTATION AND COMPUTATION-THEORETIC REALIZATION*

During the past several decades, talk of symbolic processes has become prevalent in the philosophy of mind, linguistics, and cognitive psychology. Explanations of linguistic and other cognitive capacities—including propositional thought itself—invoke symbolic entities and formal computations. While there certainly is dispute over whether and how computational concepts apply to philosophical discussion and scientific research of the mind, no one can deny that this dispute is of interest and importance.

A key question concerning this so-called 'symbolic turn' has to do with the relation between the formal processes and physical processes that are claimed to take place within our brain. (This question regarding the connection between the formal and the physical also arises in other contexts—in particular, with respect to digital computers. However, it clearly is more urgent in the discussion of the human mind, for reasons I will touch on below.) Hardly anyone who appeals to computational symbolic processes in an account of the mind holds that there are formal entities and procedures in our heads above and beyond the brain's physics. Rather, the formal is viewed as being realized by the physical: Symbolic entities and processes are embodied in the brain in a concrete, bio-chemical fashion that is inessential to their formal properties and hence also inessential to anything that might be explained (or constituted) by these formal properties. The somewhat loose connection between the two levels is supposed to be one of the major advantages of the proposed type of explanations.

* This research was supported by the Israeli Science Foundation (Grant No. 153/2004). I am grateful to Hilary Putnam for discussing with me some of the questions raised in this paper.

0022-362X/10/0706/275–292

However, this very connection—that is, the concept of realization—has been subject to substantial attack. What can justify a claim that some physical systems realize formal processes and others do not, or that a given physical system realizes a certain formal process and not another? Opponents of computational accounts of the mind, such as John Searle,[1] have given negative answers to these questions: Nothing can justify such claims without presupposing the very concept that is meant to be explained, namely, intentional cognition (which ascribes formal interpretation to some physical processes for pragmatic reasons). Philosophers of computation and of mind who disagree with Searle have rebutted these claims, arguing that the realization of the symbolic by the physical can be given solid grounding,[2] and therefore that the notion can serve the explanatory purposes with which it is burdened.

It is surprising that throughout this debate there has been little consideration of another, much older relation between formal and physical entities: the association between numbers and the physical world that is made in measurement—for example, of length or temperature. Critical examination of the foundations of measurement only began at the end of the nineteenth century, and even since then philosophers have given measurement little attention. Nevertheless, in the second half of the twentieth century, thought on measurement coalesced into the widely accepted Representational Theory of Measurement (RTM). It is therefore natural to ask whether and how the association of numbers with the physical world (according to RTM) relates to the association of symbols and symbolic processes with it: Are these one and the same? If not, why, and how are they different? It seems that answering these questions should advance the discussion of realization, whatever the answers may be.

The objective of this paper is to explore the connection between measurement-theoretic representation and computation-theoretic realization. In the first section, I give a short overview of RTM and the way it construes the connection between numbers and physical objects. In the second section, I elaborate the notion of realization. In the third section, I ask whether there are grounds for holding that realization and (measurement-theoretic) representation are distinct relations or concepts, and I answer this question negatively. I argue that it is a single relation (or concept) viewed from different angles (due to different underlying motivations). In the fourth and final section, I begin pursuing the consequences of this conclusion.

[1] Searle, *The Rediscovery of the Mind* (Cambridge: MIT, 1992), p. 209.
[2] Jack Copeland, "What is Computation?" *Synthese*, CVIII (1996): 335–59.

I. THE REPRESENTATIONAL THEORY OF MEASUREMENT

What does measurement consist in? Under what conditions can we justifiably assign numbers to objects as measures of certain properties that they have? These questions were first studied by Hermann von Helmholtz in the late nineteenth century,[3] later worked on by Otto Hölder, Norman R. Campbell, and Stanley S. Stephens,[4] among others, and given a comprehensive, formal treatment by David Krantz, Duncan Luce, Patrick Suppes, and Amos Tversky in their *Foundations of Measurement*.[5] This last work reflects a synthesis of earlier efforts,[6] and it presents the Representational Theory of Measurement (RTM) in a mature and detailed fashion.

According to RTM, measurement begins with a system of objects and some primitive relations and operations on them. For example, the measurement of length begins with the relation 'longer than' and the operation of concatenation. (The relation 'longer than' is induced by a primitive comparison procedure, in which objects are put next to each other in order to verify which goes beyond the other. The concatenation of two objects is the result of putting one next to the other in the right way.) In RTM, such a system is called an *empirical relational structure*; the adjective 'empirical' indicates that the question of whether or not a relation obtains between a given *n*-tuple of objects can be given an empirical answer, and that the operations should be executable. (What these requirements amount to is matter for a discussion not taken up here.)[7]

An empirical relational structure—in particular, its relations and operations—can have certain formal properties. For example, the relation 'longer than' can be shown to have the formal properties of a weak ordering (for example, to be transitive), and concatenation is commutative. If the empirical structure satisfies a certain set of such

[3] Helmholtz, "Numbering and Measuring from an Epistemological Viewpoint," in his *Epistemological Writings* (Dordrecht: Reidel, 1977), pp. 70–108.

[4] Hölder, "Die Axiome der Quantität und die Lehre vom Mass," *Berichte über die Verhandlungen der Königlich Sächsischen Gesellschaft der Wissenschaften zu Leipzig: Mathematische-Physikalische Klasse*, LIII (1901): 1–64; Campbell, *An Account of the Principles of Measurement and Calculation* (London: Longmans, 1928); Stevens, "On the Theory of Scales of Measurement," *Science*, CIII (1946): 667–80.

[5] Krantz et al., *Foundations of Measurement*, 3 vols. (New York: Academic Press, 1971–89).

[6] See José A. Díez, "A Hundred Years of Numbers: An Historical Introduction to Measurement Theory 1887–1990: Part I: The Formation Period: Two Lines of Research: Axiomatics and Real Morphisms, Scales and Invariance," *Studies in History and Philosophy of Science*, XXVIII (1997): 167–85.

[7] There are also second-order, so-called Platonist versions of RTM, according to which the above described relations hold among first-order properties of objects. See Brent Mundy, "The Metaphysics of Quantity," *Philosophical Studies*, LI, 1 (1987): 29–54.

requirements, it can be proved that there exists a homomorphic (structure-preserving) mapping from the structure into the numbers—typically (as in length measurement) the real numbers. That is, (1) each object is assigned a number, and the empirical relations and operations are assigned numeric relations and operations, and (2) this is done in such a way that an *n*-tuple in the empirical structure satisfies an empirical relation if and only if the numbers assigned to the objects in question satisfy the relevant numeric relation (and similarly for the operations). In the length example, the numeric relation 'bigger than' is assigned to the empirical 'longer than', and numeric addition is assigned to empirical concatenation. A theorem stating that such a mapping exists in a given case (that is, with respect to an empirical structure satisfying a given set of requirements or axioms) is called a *representation theorem*.[8]

A measurement-theoretic representation theorem typically is accompanied by a *uniqueness theorem* that states how all the homomorphisms from the given (type of) empirical structure to the numerical one relate to each other (that is, can be obtained from each other). For example, any two homomorphisms from the empirical structure that underlies length measurement can be obtained from each other through multiplication by a constant; a switch from one such homomorphism to another is just a change of scale, for example, from measurement of length in feet to metric units. Proving representation and uniqueness theorems is what measurement theory is about, according to RTM. In any given scientific case, be it physics or psychology, proving these theorems is the only required (and possible) justification for applying measurement.

Note that the proof of a representation theorem establishes the existence of a homomorphism *from* the empirical structure *to* the numerical one, not an isomorphism. (Thus, for example, this homomorphism need be neither one-one nor onto.) At the same time, the requirement on the mapping (that is, the definition of a homomorphism) is bi-directional in the following sense: If an empirical 'fact' obtains, then so does its numerical counterpart, *and vice versa* (there is an 'if and only if' in the requirement). This feature allows us to go against the direction of the mapping and infer the existence of empirical relations among objects from the existence of corresponding numerical relations among their measures. Indeed, this is why we make use of measurement. We use our

[8] In the proof of representation theorems the physical relational structures are replaced by isomorphic abstract algebraic structures (satisfying the same axioms). Thus such theorems are of a pure mathematical nature, and apply to any physical relational structure that satisfies the formal axioms involved. In accord with current literature on measurement, though, representation theorems will be presented here as relating the physical directly to the numerical. The role of the algebraic intermediary will be left implicit, except when this role bears upon the goals of this paper.

knowledge of numbers to keep track of the empirical domain to which they are assigned. (Chris Swoyer calls this practice *surrogative reasoning*.[9])

According to RTM, the existence of a homomorphism from the empirical structure to the numerical structure is all that is required for measurement, and therefore there are no theoretical grounds for preferring one such homomorphism over another. This aspect of RTM accords with our intuition that changes of scale in measurement are legitimate: Any scale is as correct as any other. (As noted above, measurement-theoretic uniqueness theorems investigate the scope of possible changes in scale in any given case of measurement.) Moreover, the aforementioned aspect of RTM (the sufficiency of a homomorphism for measurement) implies that there is even greater freedom in our choice of a numeric structure for measurement than what uniqueness theorems capture. Given a specific empirical structure, we can change not only the numbers assigned to objects, but also the numeric relations and functions assigned to empirical relations and functions. As long as a representation theorem can be proved there are no grounds for preferring one assignment over another (of course, there might be pragmatic reasons for doing so). For example, in the case of length measurement, concatenation can be assigned numeric multiplication rather than addition, and the resulting array of assignments of numbers to objects (which are yielded by exponentiation of standard assignments) will be as good as the additive version that we are used to. Thus we see that RTM allows for a loose connection between the physical and the formal—a structure-preserving mapping is enough, and no such mapping is superior to any other.

II. THE REALIZATION OF FORMAL COMPUTATIONS BY PHYSICAL PROCESSES

We now turn our attention to the domain of computation. In this domain, questions arise that are similar to those considered vis-à-vis measurement: What is required from a physical process in order for it to realize a computation? On what grounds can a certain physical process be said to realize a certain computation and not another? We saw above that the study of analogous questions concerning measurement began only in the late 1800s, centuries after the practice of measurement became so important to various human endeavors. Therefore, it should come as no surprise that similarly fundamental questions regarding the much newer theory of formal computation have received relatively little attention.

[9] Swoyer, "Structural Representation and Surrogative Reasoning," *Synthese*, LXXXVII, 3 (1991): 449–508, sections I–VI.

In this context, symbolic, formal computation should be distinguished from other, typically older notions of computation. As indicated by examples from the abacus to Babbage's Difference Engine, computing machines are not new at all. However, the (most probably only implicit) notions of computation underlying these machines need to be distinguished from the symbolic notion, initiated by the work of Turing and other logicians in the 1920s and 1930s, which has provided the basis for much work in contemporary computer science, cognitive science, and related fields. According to this view, a computation is an algorithmic manipulation of strings of symbols from a certain finite alphabet. The fundamental notion of a Turing Machine is well known for formally characterizing the previously intuitive notion of algorithmic manipulation. However, it is also important for placing the formal notions (such as a symbol, a string of symbols, and a symbolic manipulation) in the abstract, mathematical domain. Once they are so placed, it can be asked when and how they can be realized physically. By contrast, older notions of computation made less clear distinctions between the physical, formal, and numeric; hence, we shall set these aside.

Modern digital computers are commonly viewed as physically realizing algorithmic symbolic processes over a binary alphabet. Nevertheless, in scientific or practical discussions of computers there is seldom hesitation about what ensures that a given physical process (if carried out as expected) realizes a given formal one. It is taken for granted, for example, that a certain level of charge realizes (the bit) 0, that another (approximate) level realizes 1, and that certain electrical processes realize formal operations on (strings of) these digits. It is only in discussion of the human mind that questions regarding the realization of the formal by the physical come to the fore. (Interestingly, this was also the situation with respect to measurement. As long as measurement was practiced only in the physical sciences, it was not viewed as problematic and received relatively little attention. But in the early decades of the twentieth century, when quantitative measurement began to be applied in the budding science of psychology, things started to look less clear, and the question of what can and cannot be measured became a matter of dispute. The representational theory of measurement, surveyed in the previous section, can be said to be the fruit of the conceptual advances that resulted (at least in part) from this dispute.[10])

One key aspect of the realization of the formal by the physical that has received much attention and been imbued with significance is *mul-*

[10] Díez, *op. cit.*, pp. 179–84.

tiple realizability,[11] which allows for a seemingly desirable gap between the physical and cognitive. It enables us to explain, for example, how physically dissimilar creatures can still have similar or even identical cognitive states. If ideas are symbolic and cognition is computational, then my mind and the ammonia-based one of the man from Jupiter can entertain the same ideas because they realize the same formal structures through different physical means. (It should be noted, though, that the question of multiple realizability of the symbolic by the physical is distinct from, even if related to, the question of multiple realizability of the *mental* by the physical. We are concerned here only with the former.)

Discussion of multiple realizability typically circumvents the questions with which this section and the previous one began: It seldom is asked what is required for a given physical process to be a member of the family of processes that realize a given formal computation. Negative claims have been made, to the effect that no answer to this question can be given using physical concepts:[12] We should not look for any kind of physical similarity among all the possible processes that realize a certain formal process precisely because of the looseness in the connection between the formal and the physical. Even if this point is granted, however, does it necessarily follow that nothing informative can be said about the conditions which physical processes must satisfy in order to realize formal processes?

This lacuna, whether due to necessity or neglect, has opened the door for several authors to stage a major attack on computational accounts of the mind.[13] It is not only that realizability is multiple, they say. Rather, it is arbitrary: Proponents of the formal conception of mind are silent about what is required from a physical object for it to run a given computation because no such desiderata exist, and any sufficiently complex object can be said to run any formal computation whatsoever. For example, Searle gives an outline of an argument (fleshed out by Copeland) that certain regions and processes in the wall behind him, described at the atomic level, can be correlated with formal entities so as to justify the (absurd) claim that the wall runs a

[11] Ned Block, "Can the Mind Change the World?" in George Boolos, ed., *Meaning and Method: Essays in Honor of Hilary Putnam* (New York: Cambridge, 1990), pp. 137–70; Jerry Fodor, "The Mind-Body Problem," *Scientific American*, CCXLIV (1981): 114–23; Hilary Putnam, "The Nature of Mental States," in Block, ed., *Readings in the Philosophy of Psychology*, vol. 1, (Cambridge: Harvard, 1980), pp. 223–31.

[12] Block, "Can the Mind Change the World?" and Fodor, "Special Sciences, or the Disunity of Science as a Working Hypothesis," *Synthese*, XXVIII, 2 (1974): 97–115.

[13] Searle, *op. cit.*, and Putnam, *Representation and Reality* (Cambridge: MIT, 1988), appendix.

word-processing program. The conclusion of this argument and others like it is that the ascription of formal dimensions to physical processes is subject-relative and pragmatic in nature. The realization of the formal by the physical cannot serve as a cornerstone in an explanation of what cognition is: It presupposes intentional, subjective thought and hence cannot underlie it.

There have been various responses to this line of argument, and it is beyond the scope of this paper to review all of them. Nevertheless, it should be noted that any counter-argument that aims to decrease or eliminate the complete arbitrariness that Searle sees in the connection between the physical and the formal must address, even if indirectly, the questions raised here: If you want to show that realization is not arbitrary, you have to elaborate in what cases it obtains and in what cases it does not. For example, in order to block Searle's move, Copeland suggests a distinction between standard (or normal) physical realizations of a given formal procedure and nonstandard such realizations.[14] One condition that standard realizations should satisfy, according to Copeland, is that the physical pattern through which they realize the formal pattern should have modal, counterfactual-supportive force (to mirror the law-likeness of algorithmic formal steps). This condition should rule out Searle's untamed realizations: These realizations use ad hoc connections between physical loci within a given physical body and formal entities, which are good for a given computation process but would not work for any hypothetical, alternative computation (with the same architecture). Now, without assessing Copeland's suggestion, we can see that it indeed is presented as a criterion that needs to be met by a physical system in order for it to realize a formal computation.

We see, then, that considering the connection between the physical and the formal in the context of human cognition brings to the fore questions that previously have not received attention. As noted above, this was also the case with the study of measurement.

III. REPRESENTATION AND REALIZATION: TWO RELATIONS BETWEEN THE PHYSICAL AND THE FORMAL, OR A SINGLE ONE?

The previous two sections characterized both the realization of formal computations by physical processes and the measurement-theoretic representation of physical (so-called empirical) structures by numeric structures as relating the physical domain to the formal, mathematical domain. This characterization opens the door for the question whether these are indeed two distinct relations between the physical and the

[14] Copeland, "What is Computation?"

formal or the same relation arising in (or applied to) two different contexts. This is the central question of this paper.

It is clear that at first (or even second) glance, the two relations in question seem distinct. To begin with, the formal is dissimilar in the two cases: In the case of computation it consists of symbolic entities and processes, and in the case of measurement, of numbers. Second, the motivations underlying the association of the formal with the physical in the two cases can be described as opposing each other: In measurement we use numbers to keep track of the physical world, and in computation we use physical systems to yield formal results. Third, in measurement we find the phenomenon of scale-change, and it seems that in physical realization of computations we find nothing analogous. Finally, the metaphysics of representation and realization seem radically different: In the case of measurement we have a representation relation between two domains (the physical and the numeric) that are viewed as separate, and in the case of computation we talk of the physical as *realizing* the formal, that is, of physical objects and processes that themselves embody formal objects and processes. For all these reasons, the fact that these two relations are hardly ever considered together or compared with one another appears plausible and justified.

However, I argue that these appearances are misleading, and, in fact, there are no convincing grounds for distinguishing between these two relations. Rather, we have here a single relation. It appears differently in the two contexts for reasons having to do, in large part, with differences between the perspectives from which it is considered and, in smaller part, with the historical baggage carried by some of the notions involved. In the remainder of this section I shall substantiate and argue for these claims.

To begin with, the difference in the formal structures appealed to in the two cases should be acknowledged, of course, but then set aside: It is not a real obstacle to assimilating them. It is true that paradigmatic cases of measurement involve the real numbers, but around this core we find an appeal to more elaborate structures, such as vectors over numeric fields.[15] Furthermore, Louis Narens has coined the notion of *abstract measurement* to designate various kinds of homomorphic mappings of nonmathematical structures into mathematical ones that are not necessarily numeric.[16]

The conception of the relation between the physical and the mathematical as a structure-preserving mapping can be extended, then,

[15] Krantz et al., *Foundations of Measurement.*

[16] Narens, *Abstract Measurement Theory* (Cambridge: MIT, 1985).

beyond the bounds of numeric measurement. Therefore, I propose that this conception also should be applied in the case of computation—that it, that the computation-theoretic case be assimilated to the (generalized) measurement-theoretic case. In what follows, I argue that a homomorphic mapping from physical processes to formal symbolic processes suffices for all that is required and expected from what is usually called the realization of the symbolic by the physical.

It is true that the mapping from the physical to the formal serves different purposes in measurement and computation. As already noted, the roles of the physical and the formal are reversed when we switch from one case to the other, in the following sense. In measurement, the typical goal is to keep track of the physical domain, and the formal domain (and its association with the physical) is the vehicle by which this goal is achieved. In computation, formal results are the goal, and the physical domain serves as the vehicle by which this goal is realized. This pragmatic shift of orientation is important in many respects (some of which we attend to below), but it does not affect the nature of the mapping relation in the two cases, which is identical.

To appreciate the fact that the change in perspectives is only at the pragmatic level, it is helpful to note that in both cases the direction of the mapping function is the same: Its domain is the physical, and its range is the formal. In the first section, I noted that this is the case in measurement, and we easily can verify that the same holds for computation. We tag physical objects, states, and processes with formal labels,[17] creating a function from the physical to the formal. Thus, when we consider a specific physical realization of a formal computation, the same tags can be assigned to distinct physical entities (at least in the sense that the binary symbol 0 can be realized concomitantly by many physical units), but the same physical entity cannot be assigned different tags. The homomorphism, therefore, does not change direction when we move between the two cases: It is the same kind of mapping looked at from different directions because of the different interests involved.

As further support for the claim that representation and realization are one and the same, consider analog computing. With computing of this kind there is no symbolic representation of the numeric inputs and outputs of the function F, the values of which need to be computed. Rather, our knowledge of the (mathematically couched) physical laws governing the dynamics of a given physical system ensures us that if x is the measure of a certain property of the system at the beginning of computation, then $F(x)$ will be the measure of a possibly different

[17] Copeland, *op. cit.*, p. 338.

property of the system at the end of the process. Clearly, in this case we should have no problem saying that x and $F(x)$ both represent and are realized by the said properties of the system. I argue that we should have no problem saying this in the case of symbolic computation, either.

The characterization of representation and realization as involving opposing perspectives on similar mapping relations can help us deal with another point of dissimilarity between the two contexts presented above. The issue in question has to do with scale-change: Such change is possible in measurement but apparently not in the realization of computations, and this difference needs to be addressed if the two contexts are to be assimilated. However, the perspective presented above shows that in fact there is an analogue of scale-change in the computational context: It is multiple realizability. Here is why. In measurement, scale-changes are possible because a homomorphic mapping from the physical domain into the formal is all that is required. Given a fixed physical domain, we are free to assign it any homomorphic image whatsoever. In the realization of formal computations, things are the same: Any structure-preserving mapping from the physical to the formal is as good as any other, so there is freedom of choice here, also. However, in this context the formal side is the objective of the mapping and so is kept fixed: We want a physical process that can realize a given formal computation. Therefore, the freedom allowed by the weakness of the mapping requirement arises in the form of multiple realizability: Any physical process that can be mapped into the formal one is as good as any other. We see, then, that scale-change and multiple realizability turn out to be mirror images of each other.

Thus we turn to the final issue raised above. If measurement-theoretic representation and computation-theoretic realization are to be assimilated, why does one of the relations (representation) uphold the distinction between the physical and the formal while the other relation (realization) ontologically reduces the latter to the former? We do not say that a three-meter-long body is an instantiation of the number 3, but we do say that a physical body in a given state is token-identical to an instantiation of the symbol it realizes. How can such distinct metaphysical outlooks be reconciled?

If, indeed, the mapping relation between the physical and the formal suffices to describe what happens when formal computations are realized, then, I argue, the ontological implications of the term 'realization' should not carry any philosophical weight. If we can understand how we perform computations by associating physical entities and processes with formal counterparts, we need not make the claim that physical entities embody formal entities—whatever that means. We can either keep the term 'realization' or replace it, but what is important is (1) that we not read more into the term than there is to it (in this context), and (2) that it not hinder us from acknowledging that indeed in this respect too there is no difference between measurement and computation.

Further, I argue that the misleading use of 'realization' in this context stems from a conception of symbols and symbolic processes distinct from the one with which we are concerned. According to this alternative conception, symbol types are what Charles Parsons[18] calls "quasi-concrete entities"—they are abstract objects "determined by intrinsic relations to concrete objects."[19] Parsons goes on to say: "Strings and expressions are the clearest case: These are types that are instantiated by concrete objects (tokens), and what object a type is is determined by what is or would be its tokens" (*ibid.*). Thus, symbols (and strings thereof) are instantiated, or realized, by their tokens, and, indeed, these tokens are viewed as symbolic entities. (Parsons labels the relation between the quasi-concrete object and its concrete tokens a *representation* relation, which he claims is distinct from the relation that obtains between a sign and its meaning. Obviously, this relation is also distinct from measurement-theoretic representation.)

The conception of formal processes and entities that concerns us here, however, is that such entities and processes are (again, in Parsons's terminology) *pure-abstract*—as opposed to quasi-concrete—mathematical entities.[20] The symbols introduced in the context of the definition of Turing machines, for example, are not identified by their intrinsic relations with concrete, observable tokens; rather, their introduction is a matter of mathematical fiat. Indeed, both types and tokens of these symbols are abstract entities, as is the tape they are written on and the machine that manipulates them. Subsequent work on computational symbolic processes has followed Turing's lead, establishing computability theory as a legitimate, abstract mathematical field that deals with entities not dependent in any essential way on their connections with concrete, physical tokens. They are not types of such tokens, as quasi-concrete entities are, but self-standing abstract objects.

Now, I do not claim that this abstract conception of symbols and symbolic processes should be preferred in all contexts over its more traditional alternative—the two can and do coexist. Rather, the thrust of these considerations is that the association of *abstract* formal symbolic processes with physical reality, which is invoked in various

[18] Parsons, *Mathematics in Philosophy* (Ithaca: Cornell, 1983); *Mathematical Thought and Its Objects* (New York: Cambridge, 2008).

[19] Parsons, *Mathematics in Philosophy*, p. 25.

[20] See also Eli Dresner and Ofra Rechter, "Turing vs. Hilbert on Symbols," unpublished manuscript.

scientific and philosophical contexts, can and should be conceptualized along measurement-theoretic lines, that is, as involving a structure-preserving mapping from the physical domain to the mathematical one. Such conceptualization (as shown above and argued below) yields various philosophical benefits that recommend it over its more ontologically loaded alternative.

IV. IMPLICATIONS AND APPLICATIONS

The above considerations provide sufficient grounds for assimilating the cases of measurement-theoretic representation and computation-theoretic realization. In this, final section of the paper, I begin to pursue the philosophical implications of such an assimilation.

IV.A. Representation Theorems in the Computational Context. Recall the following problem, raised in the second section. On the one hand, proponents of computational models of the mind regard the fact that various physical implementations can realize a given computation as philosophically illuminating. On the other hand, we lack a clear articulation of what desiderata physical systems must satisfy in order to realize formal processes. This lack opens the door to claims that realization is practically universal (that is, any sufficiently complex physical process can be tagged with any computation whatever), and further, that ascribing a computational dimension to physical processes can have no explanatory value. So the question is this: How can we uphold the view that no physical generalization can capture all and only the realizations of a given formal symbolic process, and still formulate some requirements that physical systems must satisfy in order to realize such a process?

The suggested assimilation of the measurement-theoretic case and the computational case provides an answer to this question. In measurement theory the criteria that physical systems must satisfy in order for them to be mapped into the numbers are stated in representation theorems. Such theorems do not deal directly with physical concepts and laws. Rather, a representation theorem stipulates formal axioms that empirical relational structures must satisfy in order for it to be possible to embed them homomorphically within the real numbers. (As noted in footnote 8, a typical representation theorem proves that there is a structure-preserving mapping to the numbers from an abstract algebraic structure, which is isomorphic to the physical relational structure.) Thus, for example, the physical content of the relation 'longer than' and the length-concatenation operation is abstracted away in the representation theorem which proves that length can be measured numerically. What the theorem retains are the formal properties that these two must satisfy in order for numeric representation to go through. (The relation 'longer than' needs to be an order; concatenation must be commutative and positive with respect to this order, and so on.) Thus, restrictions are put on empirical relational structures in measurement-theoretic representation theorems without any unwarranted attempt at physical generalization: The physical relations and operations that satisfy the axioms can come from any domain whatsoever.

Computation, too, requires the formulation and proof of representation theorems, in this case specifying what formal conditions a physical system must satisfy in order for it to be mapped into (and thereby realize) a symbolic formal computation. This way no unwarranted ties are made between the formal and physical, but informative claims still can be put forward regarding which physical processes do or do not realize formal computations. Indeed, Wilfried Sieg's recent work aims at this very goal.[21]

Sieg analyzes Turing's arguments for his formal characterization of (human) computation and extracts several restrictions that any *computor* (that is, a human who computes) must satisfy. (These restrictions include boundedness, locality, and determinacy—see *op. cit.* for details.) However, Sieg is dissatisfied with the fact that it is a *thesis* that relates human computors to Turing machines. Why, he asks, can we not formally characterize the class of all human computational processes and relate it to Turing machines without any talk of theses? In order to achieve this goal, Sieg presents an *axiomatization* of human computation (as well as of mechanical computation, in which parallel computing is allowed). Following Robin Gandy,[22] he considers so-called discrete dynamical systems—classes of hereditarily finite set-theoretic structures and transformations thereof—and formulates axioms governing them which capture the above-cited restrictions. All and only dynamical processes that are isomorphic to one of the abstract structures in the class—and therefore satisfy the axioms governing the class—count as (equivalents of) human computation processes. Since it can be proved that Turing machines can carry out all computations yielded by structures in this class, it

[21] Sieg, "Mechanical Procedures and Mathematical Experience," in Alexander George, ed., *Mathematics and Mind* (New York: Oxford, 1994), pp. 71–114; "Calculations by Man and Machine: Conceptual Analysis," in Sieg, Richard Sommer, and Carolyn Talcott, eds., *Reflections on the Foundations of Mathematics* (Natick, MA: Association for Symbolic Logic, 2002), pp. 390–409; "Church without Dogma: Axioms for Computability," in S. B. Cooper, Benedikt Löwe, and Andrea Sorbi, eds., *New Computational Paradigms: Changing Conceptions of What is Computable* (New York: Springer, 2008), pp. 139–52.

[22] Gandy, "Church's Thesis and Principles for Mechanism," in Jon Barwise, H. Jerome Keisler, and Kenneth Kunen, eds., *The Kleene Symposium* (Amsterdam: North-Holland Publishing Co., 1980), pp. 123–48.

follows that all human computations are indeed within the bounds of Turing computability. Thus, no talk of theses is required. The axiomatization of the phenomenon in question (through an appeal to an abstract, set-theoretic construction) allows its direct and rigorous association with Turing's restricted class of symbolic processes.

Sieg[23] argues that this analysis of human computation realizes Gödel's hopes of characterizing the notion of computation axiomatically and thus sets it on solid grounds. (Gödel expressed such hopes out of dissatisfaction with Church's suggestion to identify effective calculability with lambda-definability.) However, the important point in the context of our discussion here is that Sieg's treatment of computability manifests the fundamentals of measurement-theoretic representation theorems, as outlined in section II. In both cases, axiom systems are viewed as applicable to nonformal reality (by way of an abstract mediation—algebraic in the case of numeric measurement, and set theoretic in the case of computation). Structures that satisfy the axioms are then proved to be correlated with more accessible and manageable (both intuitively and theoretically) mathematical constructions—the real numbers in the case of measurement, and Turing machines in the case of computation. (Sieg himself[24] talks of the correlation that he establishes as a representation theorem but uses the term in a general, mathematical sense rather than a measurement-theoretic one. However, as the structures that Sieg aims to represent ultimately are nonformal, the measurement-theoretic interpretation of the term 'representation' is also in order.)

Admittedly, the mapping that interests Sieg (from his axiomatized computors to Turing machines) is not a homomorphism. However, this difference from the measurement-theoretic case can be accommodated in two ways. First, even if the representation of computors by Turing machines is not homomorphic, it does allow for what Swoyer calls surrogative reasoning, which arguably is the essence of measurement: The study of Turing machines is brought to bear upon our understanding of human computation (and, in particular, its limitations) through this representation. Second, it is possible (and for some purposes probably also desirable) to associate a given computation not with a Turing machine (that computes the same function), but rather with an abstract symbolic process that is homomorphic to it. Such a process will be based on an architecture[25] that most probably is much richer in its primitive operations than Turing's model.

[23] Sieg, "Calculations by Man and Machine," p. 403.
[24] Sieg, "Church without Dogma," p. 141.
[25] Copeland, *op. cit.*, p. 337.

We see that Sieg's work can be characterized as realizing the measurement-theoretic framework within the domain of computation. This characterization may help enable his work to bear upon the questions, as presented in section II, with which contemporary cognitive scientists and philosophers of mind are concerned.

IV.B. The Metaphysical Status of Formal Realization versus Quantitative Measure. As noted in section II, those who argue that realization is arbitrary wish to support the claim that the assignment of formal properties to physical processes is subjective. As Searle (*op. cit.*) says, if such an argument goes through, the use of computational conceptualization to explain thought and subjectivity yields a vicious circle and therefore should be rejected. However, the foregoing analysis shows that this argument can be blocked: In the cases of both measurement and computation, although aspects of the connection between the physical and the formal are human-dependent, this in no way subtracts from the objective status of the connection's actual content.

Consider measurement first. As we saw, the choice of scale in measurement is human-dependent; even the choice of a numerical representing structure (for example, additive rather than multiplicative, in the case of length measurement) is the result of human convention. However, does this mean that statements which make use of measurement are not objective? The answer is obviously negative. What is left to human choice is the identity of the mathematical structure into which the physical one will be mapped, as well as the exact homomorphism through which the mapping will be realized. What remains objective is the identity and properties of the empirical structure in question, as well as whether a homomorphism does or does not exist between this empirical structure and any given mathematical construction. The objective content of measurement is exhausted by the information on the empirical structure provided by the numerical representation, which is the same whatever the above-mentioned human choices may be.

I argue that the computational case is the same. It is true that we can choose what physical processes to use in order to realize a given symbolic computation, and maybe also which symbolic computation to appeal to (either theoretically or practically) among those that a given physical system might concomitantly realize. However, this is not to say that the connection between the realizing physical process and the realized computation in every such case is not objective. As in the case of measurement, this connection consists in a mapping between a physical relational structure and a formal structure, and once the identity of the two structures is fixed it is not a matter of human perspective or choice whether one can be mapped into the other.

Thus, contrary to what Searle says, there is no metaphysical difference between quantitative properties of physical objects (such as length and mass) and symbolic properties of physical processes (that is, their realizing some formal process or other). In both cases we have a structure-preserving mapping from the physical to the formal.

IV.C. Computation, Cognitive Representation, and Functional Role. The previous subsection showed that the view suggested here upholds the objective status of computation against attacks such as Searle's. In this subsection, I note several points on which this view is at odds with those of Searle's opponents, who promote computational theorizing about the mind.

First, consider Jerry Fodor's slogan "no computation without representation,"[26] according to which, in order for a system to be ascribed computational status, it needs to be construed as representing information of some kind. According to the view proposed here, we should reject this slogan: A physical process can be assigned a computational process if and only if it can be mapped to this process, regardless of whether it represents any aspect of its environment. It might only be useful or interesting to describe a physical system in computational terms in cases where it is construed as processing information, but this does not bear on whether it is possible for the physical system to be mapped to the formally defined computation (thereby realizing it). Thus, the view advocated here relates computation to representation in the measurement-theoretic sense and frees it from conceptual dependence on representation in the cognitive sense. The measurement-theoretic relation might underlie the representation of information, but it does not depend on it.

Further, whether a certain physical system performs computation does not depend on its functional role within a wider context. The fact that it computes a certain function (that is, is homomorphic to a certain symbolic process) may allow it to fulfill a certain function, but the term 'function' is used here in two different senses, on different levels, and the former does not depend on the latter.

Finally, it is misleading to talk about the computational identity, or individuation, of physical processes.[27] Such talk is obviously related to the notion of realization (of the formal by the physical), but once we acknowledge that, in this context, realization amounts to mapping, it becomes apparent that identity is not an issue here. A given physical system clearly can be assigned various numeric values indicating its measure in various respects, and no question arises concerning the

[26] Fodor, "The Mind-Body Problem."

[27] See for example Oron Shagrir, "Why We View the Brain as a Computer," *Synthese*, CLIII (2006): 393–416.

numeric identity of the system. Similarly, it might be that a system can be mapped onto various computational processes, but this need not (and should not) give rise to the question of which of them it is identical with. Oron Shagrir[28] may be right that semantic considerations guide us in deciding which (if any) symbolic processes should be assigned to a physical system in order to describe its dynamics. However, this does not imply that the semantically relevant assignment is somehow identical with the physical process mapped into it.

ELI DRESNER

Tel-Aviv University

[28] Shagrir, *op. cit.*, and "Content, Computation, and Externalism," *Mind*, CX (2001): 369–400.

BELIEF *DE RE*, KNOWING WHO, AND SINGULAR THOUGHT*

In his discussion of Bertrand Russell's notion of knowledge by acquaintance, H. L. A. Hart criticizes it as a "part of the *damnosa hereditas* left by Plato to his philosophical posterity."[1] Hart was insightful and, I believe, correct to align Russell's account of knowledge by acquaintance with Plato. He presumably had in mind Plato's account of knowledge as a "grasping of being."[2] This severe judgment passed by Hart extends to the very idea of knowledge of the sort Russell envisaged, and not merely to the limited extent to which Russell thought such knowledge might be available to us. In that disagreement, I side with Hart against Russell. However, it must be said that our view is a minority view. I shall argue that the more contemporary notions of knowledge *de re* and belief *de re* are also part of that "cursed inheritance." I shall try to show that it is crucial to the notion of knowledge by acquaintance, both in Russell's version and its more contemporary versions, that it come with a cluster of epistemic guarantees. These guarantees, I shall suggest, serve to distinguish belief *de re* from the more narrowly semantic notion of singular thought. I claim that nothing answers to the concept of knowledge by acquaintance nor of belief *de re* precisely because no belief about the world comes with such epistemic guarantees. However, I shall argue that there are indeed singular thoughts. I shall also try to show that these two notions both need to be separated from the notion of knowing who, which is often used as a link between them. Moreover, we will need to explore the differences between belief states and their attributions, since I shall argue that there are conditions under which *de re* attributions of belief are true and appropriate even if there are no beliefs *de re*. Separating these often conflated notions allows us to understand the limited extent to which we can expect epistemic consequences from our semantic capacities and, in particular, our ability to think of objects.

I. RUSSELLIAN BACKGROUND

When Russell addressed the problem of our connection to the world, it seemed to him that this connection was established, not by our

*I must thank Aislinn Batstone, Adam Dickerson, Gil Harman, Ralph Kennedy, Aubrey Townsend, and, sadly belatedly, David Lewis for help relating to this paper.

[1] H. L. A. Hart, "Is There Knowledge by Acquaintance?" *Proceedings of the Aristotelian Society, Supplementary Volumes*, XXIII (1949): 69–90.

[2] See the passages in the *Theaetetus* 186^{e} ff.

0022-362X/10/0706/293–310

acting in the world, but rather cognitively, by our thinking of or judging about objects. It also seems fair to say that Russell shaped his semantical views with a weather eye on overcoming skeptical doubts which might be raised about our cognitive capacities.[3] He answered the problem by distinguishing two kinds of knowledge, knowledge by acquaintance and knowledge by description. Knowledge by acquaintance came with two important epistemic guarantees: first, that the object of the thought exists and, second, that two judgments about an object of acquaintance can be recognized as being about the same object. This doctrine at one fell swoop allows for a substantive cognitive relation to the world—we can think of objects—and answers at least some of the pressing skeptical doubts. This doctrine was supplanted by the doctrine of knowledge by description which uses Russell's more famous account of definite descriptions. Russell thought this account would enable him to construct surrogates for propositions about objects with which he was not acquainted. Russell explicitly aligned his notion of knowledge by acquaintance with singular propositions and made the point that knowledge by acquaintance was at the base of our semantic capacities. According to Russell, at the root of each case of significant 'knowing that' was a case of acquaintance or 'knowing who'. As he put it, such knowledge "brings the object itself before the mind."[4]

Here is Russell introducing the topic:

> The object...is to consider what it is that we know in cases where we know propositions about 'the so-and-so' without knowing who or what the so-and-so is. For example, I know that the candidate who gets most votes will be elected, though I do not know who is the candidate who will get the most votes.[5]

> ...It would seem that, when we make a statement about something only known by description, we often intend to make our statement, not in the form involving the description, but about the actual thing described. That is to say, when we say anything about Bismark, we should like, if we could, ...to make the judgment of which he himself is a constituent. (218)

[3] Unlike Frege, who seemed to have no interest in using his account of meaning to answer skepticism.

[4] See the letter by Russell to Frege of 12 December 1904: "I believe that in spite of all its snowfields Mont Blanc itself is a component part of what is asserted." This letter is excerpted in Nathan Salmon and Scott Soames, eds., *Propositions and Attitudes* (New York: Oxford, 1988), p. 57.

[5] Russell, "Knowledge by Acquaintance and Knowledge by Description," in *Mysticism and Logic: And Other Essays* (London: Allen and Unwin, 1918), p. 209. Page references are to this paper in this volume.

> ...in judging, the actual objects concerning which we judge, rather than any supposed purely mental entities, are constituents of the complex which is the judgment. (222)

It is tempting to say that here Russell has confused the object of the judgment with the judgment itself. However, behind this identification lies Russell's

> ...fundamental epistemological principle in the analysis of propositions containing descriptions...: *Every proposition which we can understand must be composed wholly of constituents with which we are acquainted.* (219)

Russell characterizes his principle as epistemological, and it certainly is constrained by his views about the objects of our acquaintance and our epistemic access to them. However, the principle is stated and used, not as an epistemological principle, but rather as if it were really a semantic principle, a principle about what the content of judgments and sentences we use meaningfully can be. That is, the principle concerns which propositions we can understand; it is not centrally about what we can know, though it will certainly have consequences for this. It limits our suppositions no less than our knowledge.

This epistemic constraint on our semantic capacities gives Russell his response to the skeptical challenge and fits with the broad sweep of his concerns. However, the argument Russell gives for his principle—the only argument as far as I can tell—is that

> it seems scarcely possible to believe that we can make a judgment or entertain a supposition without knowing what it is that we are judging or supposing about. (219)

Now this seems right, but it seems almost in danger of being trivial if it is not false. What can he mean by "our knowing what we are judging or supposing about"? Suppose I am making a judgment about Julius Caesar and someone asks me who I am judging about. Can't I simply reply 'Julius Caesar'? They may want more, but why should I always be able to give more? There seems to be no real reason. It might be that the judgment I made expresses all my beliefs about Julius Caesar. The question can be asked, but unless I repeat myself, I cannot answer. It is interesting that here Russell makes his account of singular thought depend on the issue of knowing who. I cannot even make a supposition about someone unless I know who I am supposing about. Russell seems to have imported his notion of acquaintance into his argument for this very notion at this point. He seems to have thought that you could always use the demonstrative 'this' to indicate what it is that you are supposing about, in other words, that you are immediately acquainted with the objects of your supposition or judgment. If I am unable to use a demonstrative in that way to indicate Julius Caesar, it would follow that I am not acquainted with Julius Caesar and that 'Julius Caesar', as I use it, is not a logically proper name. Thus I would not be able to think about Julius Caesar. I cannot have singular thoughts about him. But this is unsatisfactory.[6] There is no really good reason to think that the class of thoughts one can have about particular objects is as limited as Russell thought.

Let us then stipulate that by 'singular thought' we shall mean a thought which requires a particular object to have a property for it to be true.[7] The questions we face are: are singular thoughts limited to objects with which we are acquainted? Can I have singular thoughts without knowing who I am thinking of?

The claim that Russell makes is that acquaintance comes with very strict epistemic guarantees. These guarantees serve to distinguish belief and knowledge involving acquaintance from those involving singular thought as I have characterized it. There are two epistemic consequences of the way Russell thought acquaintance "brings the object itself before the mind." We are guaranteed, first, that the object exists and, second, that we cannot confuse the object with another which is similarly brought before our mind. Whenever we are thinking about something, it is guaranteed to exist. It would follow very readily from this, given certain considerations Russell raises about what we can be certain of, that the only things we can really think about are ourselves, our sense data, and universals.

> We have acquaintance with sense-data, with many universals, and possibly with ourselves, but not with physical objects or other minds. We have *descriptive* knowledge of an object when we know that it is *the* object having some property or properties with which we are acquainted; that is to say, when we know that the property or properties in question belong to one object and no more, we are said to have knowledge of that one object by description, whether or not we are acquainted with the object. (231)

There is a peculiarity in this view of Russell's. According to him, we can think about universals in a way in which we cannot think about

[6] *Contra* Gareth Evans in John McDowell, ed., *The Varieties of Reference* (New York: Oxford, 1982).

[7] This way of characterizing singular thoughts allows us to see that the familiar technique of treating propositions as sets of worlds would not enable us to identify singular thoughts, since two thoughts may determine the same set of worlds while only one of them is a singular thought. Consider replacing 'Julius Caesar' with a rigidified description such as 'The actual referent of "Julius Caesar"' in some simple sentence. Only the use of the name yields a singular thought. A similar point can be made using what Saul Kripke calls "de facto rigid descriptions" such as 'the smallest prime'. See Kripke's *Naming and Necessity* (Cambridge: Harvard, 1980), p. 21, footnote 21.

the objects having the universals. A universal apart from the object it inheres in is an odd thing to have before your mind. If we cannot be acquainted with physical objects, how is it that we can be acquainted with their properties? Recall that it is the universal itself that is supposed to be before your mind, not an idea of it or anything mental. Perhaps Russell meant to restrict the sorts of universals involved here to those which inhere in the objects with which you are acquainted, yourself, and your sense data. But this is not clear from what he says and anyway would either lead us to a curious view as we construct other universals (such as the universal had by all and only the kings) out of properties of sense data, or it would falsify what Russell says when he says

> We have *descriptive* knowledge of an object when we know that it is *the* object having some property or properties with which we are acquainted. (*Ibid.*)

Thus, according to Russell, the very same property is such that we can be acquainted with it and it can be had by objects (presumably even physical objects) with which we cannot be acquainted. This is a very hard condition to satisfy. It looks as though one of these conditions must be sacrificed. After all, if we can be acquainted with an object's properties, what more is needed before we are acquainted with the object itself?

The second epistemic consequence of the doctrine that acquaintance "brings the object itself before your mind" is that you cannot be mistaken about the identity of those things with which you are acquainted. If you are acquainted with x and you are acquainted with y, you thereby know whether x is the same thing as y. This follows from the fact that it is the object itself and not some idea, presentation, or aspect of it which is before the mind.

Russell can be summarized as saying that we can only have singular thoughts about things with which we are acquainted and that we are acquainted with things only when we stand in a certain privileged epistemic position with respect to those objects. When we are in that privileged epistemic position we have the object itself and not some presentation of it before our mind.

II. PERSPECTIVAL AND NONPERSPECTIVAL BELIEF

It is typical of our beliefs that we can have distinct beliefs about the very same thing and not realize that we are judging about the one object. One sort of situation in which this occurs is when we see the same thing from different perspectives. In that case we can have inconsistent beliefs, not know that our beliefs are inconsistent, and, crucially, not be able to know without more empirical information.

Let us say that beliefs which are liable to this sort of contradiction which is not discoverable without additional empirical information are perspectival. It is a consequence of Russell's account of acquaintance that states based on acquaintance are nonperspectival: such states are not liable to that sort of mistake.

Russell's view is at least consistent, if not very plausible. It seems, as Kripke has suggested, that singular propositions are far more extensively available than Russell's notion of acquaintance would allow.[8] Nevertheless, some sort of notion of acquaintance seems attractive. But once we start to weaken the idea of acquaintance from the strong epistemic constraint that Russell invoked, we lose the idea that in using a sentence which expresses a singular thought we have gained any particular epistemic guarantee. For we can use such sentences and fail to refer to anything. The epistemic guarantees that Russell alleged flow from his notion of knowledge by acquaintance are overblown and cannot be realized.

Emphasizing this perspectival aspect of our beliefs, while holding that the objects of the beliefs are not perspectival, facilitates an accommodation between a Fregean approach to beliefs and a direct reference theory of the semantics of names. Key to that accommodation is the emphasis on the difference between a semantics of beliefs and the semantics of belief attribution. It seems plain that such an accommodation is possible since the analogous case of perception is most naturally treated in precisely this way.

The natural view of perception is that we see objects, yet we see objects from perspectives and by dint of real causal relationships in which we stand. Some criticisms of the theory of direct perception, such as those made by Brand Blanshard in his review of Donald Cary Williams's *The Principles of Empirical Realism*, show that some explanation of this position is necessary.[9] Blanshard asks, for example, "How can one's apprehension of a flash [of a supernova] be called 'direct' if it is the end result of a causal process that began a century ago?"[10] This is a good question but one which can be answered by distinguishing the object of the perception, what it is of, from the process which brings it about that the perception has that object. A theory which assigns objects to perceptions need not give the ground of that assignment. However, a theory which explains the ground of the assignment of objects as objects of a perception should, given the appropriate inputs, allow the derivation of the objects which are so assigned.

[8] Kripke, *op. cit.* See in particular the discussion in the Preface.

[9] Blanshard, untitled review, *The Philosophical Review*, LXXVIII, 3 (July 1969): 399–402.

[10] Blanshard, *op. cit.*, p. 400.

It is plain then that we do indeed see objects and not some sort of perspectival aspect. We see things, the objects, from perspectives. The very thing we see from one side we can see from the other side. Insofar as we seek an object to assign as the object of perception we reach for the same object whether we see it from one perspective or another.

We know that sometimes we do not recognize that an object is the object of two distinct perceptions. And sometimes this is not a failure of reasoning ability. Sometimes the world conspires to show us the same object from two perspectives. We then can have inconsistent appearances, but appearances whose inconsistency is not discoverable as a matter of reasoning alone. This would not be possible if our perceptions had different perspectival objects, for then the appearances would not be inconsistent.

By analogy, the natural thought is that we can have multiple thoughts about the same object, even though we cannot ascertain that by reasoning alone. If those are singular thoughts then we should deny that singular thought has the epistemic properties which are associated with knowledge by acquaintance.

This natural thought is closely connected with Frege's familiar motivation for the notion of sense. It is because we can be perfectly rational and find ourselves with two beliefs which turn out to be about the same thing where more empirical information was required to find that out that Frege suggested we needed to introduce a notion of sense. Sense is that aspect of meaning which is constrained by perfect rationality and, hence, *a priori* relations.

Sense might legitimately be called an aspect of meaning, but to say this is rather ambiguous. Usually a theory which posits intermediary objects between the linguistic expressions and their referents, as does Frege's, has been taken to be inconsistent with a direct realist understanding of the semantics of proper names. However, following Kripke, we can distinguish a theory of meaning from a theory of reference.[11] Some theoretical notion like Fregean sense can have a role in a theory of meaning even if not in a theory of reference. A warning: calling it a "theory of meaning" might be misleading; nothing I have said suggests that this notion of sense is language-wide, nor even that it is constant for each speaker.[12] However, this tells us that where semantic properties of the whole depend only on the referents of the parts, this amalgam

[11] Kripke, *op. cit.* See for example, pp. 5, 32, and 53ff.

[12] In fact it is unlikely that a notion of meaning which holds these constant will be available. However, since even the epistemic status of an interpreted sentence like '*s* is one meter long' can differ across agents and contexts, this is to be expected.

theory will produce the same predictions as the view that names only contribute their referent to the semantic value of the whole.

III. BURGE'S BELIEF *DE RE* AND KAPLAN'S SINGULAR PROPOSITIONS

Russell's notion of knowledge by acquaintance was revived in recent work as belief *de re* by Tyler Burge.[13] I shall suggest that the very same problems we found with Russell here recur. According to Burge, *de dicto* beliefs are fully conceptualized[14] whereas *de re* beliefs are those "whose correct ascription places the believer in an appropriate non-conceptual, contextual relation to objects the belief is about."[15] Again, the idea is that there is a type of belief which is nonconceptual and involves the object itself. The first thing to do is to distinguish the *de re* attribution of belief from *de re* belief as such. I shall show that there is a very widespread practice of using *de re* attributions, far beyond the sorts of cases where the rich connection with the object is plausible. In that case, the idea that belief *de re* necessarily involves this sort of nonconceptual contextual relation to the object of belief can only be sustained if we are willing to forego the connection between belief attributions which are *de re* and the attitude Burge has raised. As it is, we may as well say that there are no beliefs *de re*, that all there is is the peculiar kind of attribution. This is the conclusion for which I shall be arguing. To get to that conclusion I shall first try to show that *de re* attributions of belief are widespread and play a particular role for us. I shall then try to show that that role does not involve anything like a distinctive attitude. I shall then try to show that attributions *de re* are independent of the issues of knowing who and singular thought.

When is it appropriate to say of a person that they know who they are talking about? My own view is that there is no uniquely good answer to that question and that the significance attributed to the notion of knowing who is overstated. We use different standards for different contexts; exams, coronial inquiries, and office gossip all involve different standards for 'knowing who'.

Behind the philosophical discussion of knowing who is the mistake of trying to delineate a class of beliefs which anchor us to the world because they involve the objects themselves. Russell makes it clear that this epistemic result was one of his motivations.[16] This mistake

[13] Tyler Burge, "Belief *De Re*," this JOURNAL, LXXIV, 6 (June 1977): 338–62.
[14] *Ibid.*, p. 345.
[15] *Ibid.*, p. 346.
[16] Gareth Evans (*op. cit.*) revived something like Russell's epistemic conception of acquaintance; see especially chapter 6, section VI and chapter 7, section II. However, the account Evans gives of mock thoughts and the fact that they are introspectively indistinguishable from genuine thoughts ought to give us reason to doubt the significance of the immunity from error he thinks he has discovered.

recurs in Burge's notion of belief *de re*: an attitude that involves the object itself and so avoids mistakes about its object and its existence. It is also found in David Kaplan's early work on the notion of a name which represents an object to an agent. He introduces it as follows:

> *a represents x to* Ralph (symbolized '**R**(*a*, *x*, Ralph)') if and only if (i) *a* denotes *x*, (ii) *a* is a name *of x* for Ralph, and (iii) *a* is (sufficiently) vivid.[17]

Such a name will, because it is vivid, have certain rich internal connections for the user of the name. According to Kaplan, only when we have a vivid name at the ready can we express singular propositions. Thus, singular propositions are only available to a language user for object-name pairs which have an appropriately rich set of internal connections for that user. Singular propositions are simply about the thing in question, unmediated by any conceptualization. The temptation is to conceive of these propositions in Russell's way, as containing nothing analogous to Fregean senses but rather containing the objects themselves. Similarly, the objectual and nonconceptual aspect of belief *de re* is built into it. The question raised earlier about Russell arises here with just the same pertinence. Is it possible for someone to be acquainted with an object by way of two names of the object which are each vivid and not realize that they are names of the same thing?[18] Clearly, that is a possibility. But then the contents assigned to the sentences to which the speaker will assent in using these names will not capture the entailments among this speaker's beliefs, for these will simply have the object with two properties. It is not a rational failing that the speaker does not realize that these two names are names of the same thing. Yet, given the way the propositions are assigned, there is nothing to mark the lack of *a priori* relation between the vehicles of the propositions.

There are two ways to go forward which might commend themselves at this point. The first is to require that there exist a richer set of connections between the agent, the name, and the object referred to by the name before a singular thought about that object can be attributed to that agent via that name. The thought might be

[17] David Kaplan, "Quantifying In," in Leonard Linsky, ed., *Reference and Modality*, (London: Oxford, 1971), pp. 112–44, at p. 128. [Originally published in *Synthese*, XIX, 1/2 (December 1968): 178–214.]

[18] And not to lack any relevant linguistic knowledge. Of course, we could so weight the matter that any such situation reflects some sort of inadequacy of linguistic understanding, but this would be *ad hoc*.

that given enough information associated with two names this sort of problem case will be ruled out. And it seems true that if we put strong enough constraints on singular thoughts in this way such problem cases are less likely to arise. However, the objects about which we could have singular thoughts would be similarly restricted. They would seem to have to be limited to the sorts of occurrent momentary objects with which Russell thought we are acquainted.

The second way is to give up the epistemic guarantees associated with singular thoughts and recognize that the richness of the connections associated by a speaker with a name is a matter of degree. In this way, a singular proposition can be assigned to someone who has names for the objects of the proposition whether or not they have a vivid name for the object. The social character of language ensures that we can use terms with full semantic capacity even when our associated connections with the objects or properties so named are rather tenuous.

A sentence which expresses a singular proposition can be used by someone who is not in a position to know much about who the sentence is about. Having listened often enough to sports shows on the radio, I have heard tell of Valentino Rossi, the Italian motorcycle racer. I have acquired a new name in my lexicon by a means as normal as any we encounter. If I say 'Valentino Rossi won a race' I use this sentence to express a singular proposition about Valentino Rossi. The sentence in my mouth has the usual meaning, even though the only connection I have to Valentino Rossi is that I heard this on the radio the other day. I may, but need not, know more. Perhaps the radio report details injury concerns or some such. I have no idea what Valentino Rossi looks like, though I would expect him to be a rather small person as are many professional motorcycle racers. I can be said to believe of Valentino Rossi that he won a motorcycle race. That is, exportation seems appropriate in this case, even when the richness of the information I have about Valentino Rossi, and so the vividness of the name, is rather thin. It might well be that he is also the only rider on the motorcycle tour with the initials 'V. R.'. Someone who knew that (not me!) could truly say about me that I believed of the only rider on the motorcycle tour with the initials 'V. R.' that he won a race. That belief attribution *de re* is perfectly true, even though I have not the faintest belief either way whether Valentino Rossi is the only rider on the motorcycle tour with the initials 'V. R.', and even though it would be wrong to think that I am acquainted with him or know who he is. Certainly, I may believe that he is the only rider on the motorcycle tour with the name 'Valentino Rossi', but apart from that description there is not much I know which would serve to individuate him from other motorcycle racers.[19] Thus it is wrong to think that before one can use a belief attribution *de re* the believer must have an ability to distinguish the object of the belief from all others.

In making the *de re* attribution to me the attributer is committing himself to the claim that Valentino Rossi is the only rider on the motorcycle tour with the initials 'V. R.' but is not attributing that belief to me. He presents my belief as being about the person with that property without committing himself to how I think of that person. Indeed, it seems that that is the most strikingly useful aspect of *de re* attributions: they tell us who the belief is about without telling us how that person is thought of. That is, the attribution abstracts from that aspect of the belief content. *De re* attributions do not attribute a strange unconceptualized belief; rather, they are a way of abstracting from that part of the content.

Nevertheless, the belief that I do have, that Valentino Rossi won a motorcycle race, is a singular thought. It is true if and only if that particular person has the property of winning a motorcycle race. I might have had other beliefs which were not singular beliefs involving that person. Suppose my information had been gained by my hearing that the only racer on the motorcycle tour with the initials 'V. R.' had won a race. Hearing this and being particularly trusting of the media, I could come to believe it. That belief is not a singular thought. Yet, it is still acceptable for someone to say about me that I believe of Valentino Rossi that he won a race, even though I may have not heard the name 'Valentino Rossi'. There is no need to limit the *de re* attribution to believers whose beliefs express singular propositions. We do not ordinarily do so. That being the case, there cannot be a close connection between the appropriateness of a *de re* attribution and the content of a belief.

Moreover, as we have seen, *de re* attribution is largely independent of the issue of knowing who. The mistake of conflating the two seems to have been imported into this area by Russell when he explicitly aligned the notion of knowing who with singular propositions. The purported cognitive significance of the distinction between beliefs indexed by singular propositions and those indexed by nonsingular propositions is an illusion. And in neither case can we be sure that the objects of the beliefs exist.

In the present context we can understand what might lead to the mistake: the fact that the distinctions between the many dimensions of

[19] It would be surprising to find out that there are two motorcycle racers with the name 'Valentino Rossi' but not violently surprising. It is akin to finding out there are others who share your name.

content have not been adequately distinguished. The practice we have been worrying about—the indexing of beliefs by propositions—is on the whole too simple. Beliefs have many dimensions of content. One dimension, which might be called their metaphysical dimension, is the way the world must be for what was expressed to be true. Another dimension is given by a notion of narrow content, say, the ways the world could be in which creatures with a course of history introspectibly like those of an actual person speak truly when they utter the sentence in question. I shall argue that yet another dimension is given by epistemic consequence relations.

Given this complexity we obviously have to think hard about the work belief-attribution sentences are doing. If belief is simply a relation to a proposition, and a proposition is just a way the world is or could be, then any equivalent way of reporting that way the world is should be equally acceptable as a belief attribution. But clearly this is not always the case. Failure of substitution of co-referential names or necessarily co-extensive predicates is manifestly a feature of belief attributions. So belief attributions are not merely reporting the relation of the believer to a way the world could be. Some might ask whether this something more that belief ascriptions are performing is to be counted as part of their semantics or their pragmatics. After all, might it not be that the feature of failure of substitution is to be explained by conventional implicature, a noncancellable feature, part of meaning but not part of the truth conditions? I am in general skeptical about the presence of a conventional implicature which is taken to distinguish the attribution of truth from the appropriateness of utterance, when that implicature cannot be cancelled. For, this is a form of pragmatic implicature which is invariant upon the conditions to which pragmatic factors are sensitive, namely, contexts of utterance. It is hard to imagine that we could come to predict the presence of such pragmatic factors.[20]

[20] An example: 'and' and 'but' have been presented as having the same truth conditions though differing in conventional implicature. See Paul Grice's views on the matter as presented in revised form in *Studies in the Way of Words* (Cambridge: Harvard, 1989). Theories as to the nature of the difference between 'and' and 'but', which nevertheless assign them the same truth conditions, must give an account of the way these particles behave differently. Such accounts may make use of the notion of a presupposition. It may be that a use of 'but' presupposes a contrast between the subsentences that a use of 'and' may not. However, this notion of presupposition is far from clear. It is not obvious that such a contrast even if it is presupposed is not also asserted. And for that reason it is not obvious that presupposition failure leads to a lack of meaning, or truth with the flouting of a convention, rather than simple falsity. However, it would seem that the view that 'but' differs from 'and' in presupposition but not in truth conditions is committed to an account of presupposition which depends upon what is being presupposed not being also asserted.

So what other work could a belief attribution be doing? In the discussion so far, the epistemic dimension of the content has been omitted. But that is the dimension of content which is relevant to deliberation, since it is precisely that dimension of content which is accessible to the agent. We hold people responsible, to greater or lesser degrees, for failing to recognize the *a priori* consequences of their beliefs. If belief attribution is indicating more than the metaphysical content of the belief, by indicating the epistemic content as well, we have an explanation of why belief attributions do not admit of substitution of co-referential names in general. In such an account, the semantic story offered by the direct reference theorists for sentences involving proper names can be endorsed without thereby committing to a substitution principle for names in all contexts. Thus the metaphysical content of a belief can be distinguished from the impact the belief makes on the epistemic state of the believer as a whole. In other words, the metaphysical content of a belief can be distinguished from its epistemic content.[21]

What lesson do I suggest we learn from the independence of issues surrounding the *de re* attribution of belief, belief *de re*, singular belief, and being *en rapport* or acquainted with an object from the issue of knowing who? In summary, there are no beliefs which deserve to be called beliefs *de re*. We do not attribute *de re* beliefs; rather, we use *de re* attributions of belief.

So what is it to be *en rapport* with an object? It is to stand in certain actual relations with the object, to be able to display a certain degree of practical recognitional expertise. But one thing is not involved: it should be no part of the notion that you can always recognize the object with which you are *en rapport*. This is easily seen in the case of numbers.[22] I may well be *en rapport* with the numbers less than 100. But there are many sentences about particular numbers in that set, such that I do not know which number the sentence is about. This is also quite obviously the case with many everyday objects. Being *en rapport* with my pen does not mean that I can distinguish it from all other pens. Just as being able to identify a species of animal, for example,

[21] At this point I take myself to be building on the lessons Kripke has already made (*op. cit.*). In particular, his remarks on the distinction between the necessary and the *a priori* makes, I think, a related point. We can treat the metaphysical content of a belief as its metaphysical consequences and the epistemic content of the belief as its *a priori* consequences. These consequences themselves can be taken to be sets of belief states. In this way the belief that Hesperus is a planet has the very same metaphysical but not epistemic content as the belief that Phosphorus is a planet.

[22] My interest in the topic of *de re* attitudes was sparked by Kripke's remarks in lectures at Princeton on *de re* attitudes to numbers in 1990, although he seems in favor of *de re* attitudes and it is clear I am against them.

a Pacific Black Duck, does not mean that you can distinguish it from all possible species, or even all actual species. The criteria we use for identifying species are often quite locality specific. A good criterion in one locale might turn out to be inappropriate in a new region where a number of species share some feature which was diagnostic of just one species.

This is directly analogous to the Masked Man paradox. This paradox, or apparent paradox, starts with the supposition that I know who my father is but not who the masked man is. Therefore, I infer that my father is not the masked man. But when the mask is removed it turns out that the masked man is my father. Each of the premises seems true; the conclusion then seems to follow by Leibniz's law; so what goes wrong? Or consider again the mundane example discussed by Russell: I know who each of the candidates is for the election which has taken place. There is a winner, but it is not true that I know who the winner is. Yet, one of the candidates I know is the winner. But how can that be, if I know and am acquainted with each of them already?

This paradox is an interesting puzzle that shows up a vagueness in our use of the term 'know'. If you think that knowing who involves an ability to recognize in all contexts you should say that you do not know who your father is, for here is a context in which you do not recognize your father. And if you think knowing who is a matter of acquaintance, and that acquaintance comes with the Russellian guarantees, that will lead to the conclusion that you are not acquainted with your father. Thus (albeit implausibly), you deny the first premise.

However, if you do think you know who your father is then you do not think that knowing who involves an ability to recognize in all contexts. Thus the first premise is true on this understanding, as seems appropriate. What about the second premise? On this understanding the second premise is false, for you do know who this person is, since you know who your father is. Although it seems appropriate to say that you do not know who the masked man is, this is a different sense of 'know'. Thus our understanding of the premises in the argument allows that they are each true, and true together, but not on the same understanding of the key term, 'know', involved. Therefore, it is the appeal in the argument to Leibniz's law that is illicit, not for the bad reason that this law is less than universally true, but rather because of a fallacy of equivocation.

This issue of knowing who has an interesting analogue in understanding our language. We can define a sentence '##' to mean 'Ponting is a happy cricketer' if at the moment of defining Ponting is a happy cricketer and to mean 'It is not the case that Ponting is a happy cricketer' if at the moment of defining it is not true that

Ponting is a happy cricketer. At the moment of defining we are in a position to know that ##, and we are in a position to know that *a priori*. Yet what '##' says is contingent; if Ponting is a happy cricketer then it is quite possible that he should not be, and vice versa. '##' is an example of a contingent *a priori* truth. On the other hand, while treating '##' as a sentence we know to be true seems fine at the metalinguistic level, it does not at the material level. Just which proposition does '##' express?[23]

The case of the election winner is analogous to the case of '##'. In these cases there is a range of objects and a range of propositions, with each of which you are acquainted. The question is whether in either case or both you know who the election winner is, or which proposition '##' expresses.

IV. KAPLAN'S RETREAT

I suggested that Kaplan's "Quantifying In" represents a good example of the same mistake that Russell made in restricting singular thought to objects of our acquaintance, or as Kaplan put it, to objects with which we are *en rapport* via vivid names. Kaplan saw this was wrong and quickly moved on from the views in that paper. In a later work he says:

> All this familiarity with demonstratives has led me to believe that I was mistaken in "Quantifying In" in thinking that the most fundamental cases of what I might now describe as a person having a propositional attitude (believing, asserting, etc.) toward a singular proposition required that person to be *en rapport* with the subject of the proposition. It is now clear that I can assert *of* the first child to be born in the twenty-first century that *he* will be bald, simply by assertively uttering,
>
> (29) Dthat['the first child to be born in the twenty-first century'] will be bald.
>
> I do not now see exactly how the requirement of being *en rapport* with the subject of a singular proposition fits in. Are there two different sorts of singular propositions? or are there just two different ways to know them?[24]

Do these changes mean that he moved to a position which is not vulnerable to the criticisms laid at his door? Certainly it seems that one thing he says, namely, that there is no need to be *en rapport* with the

[23] Notice that ## though true for the context of stipulation and knowable for the stipulator in that context *a priori*, might well be false in other contexts. This is analogous to the context sensitivity of truth of many of the classic cases of contingent *a priori* truths, such as the claim that the standard meter rule is a meter long.

[24] David Kaplan, "Dthat" in Peter A. French, Theodore E. Uehling, Jr., and Howard K. Wettstein, eds., *Contemporary Perspectives in the Philosophy of Language*, (Minneapolis: Minnesota UP, 1979), pp. 383–400, at p. 397.

subject of a singular proposition before being able to take up propositional attitudes to that proposition, is correct. So the positive aspect of the doctrine in "Quantifying In" is false. Being *en rapport* with the subject of a singular proposition is no part of the precondition for appropriate semantic relations to that proposition.

Since being acquainted or *en rapport* with the object of a thought is not necessary for singular thought, we are reminded again that the semantic role Russell carved out for his knowledge by acquaintance is misplaced. Even though the remarks I utter and beliefs I have about David Kaplan depend on how things are with him for their truth, I am in no special position to know that David Kaplan exists. For all I know, my trusting use of that name is like those who were fooled by the practice emanating out of Canberra a little while ago of thanking a certain Bruce Toohey for help with streamlining proofs and discussions in logic papers. This was an elaborate playful hoax played by logicians on their readers. Poor trusting soul that I am, how am I to know that the case of David Kaplan is any different? And yet, if this is not so, if the world is as I think it is, I am talking about David Kaplan. However things are with David Kaplan, the *a priori* consequences of my beliefs would be the same. On the other hand, the metaphysical content of my beliefs would change radically if I were situated in a different sort of world, but one which presented itself to me in much the way that this one has up to now.

Should we then be encouraged to endorse the second of the options offered to us by Kaplan? Should we say that there are singular propositions and "two different ways of knowing them," depending on whether you are *en rapport* with the object of the singular proposition or not? We could make that distinction, but why stop at two? There are many different ways to know singular propositions, one for each day of the week and more. The question is whether making such a distinction between ways of knowing singular propositions actually plays any theoretical role for us. The distinction Kaplan suggests will not play any significant role if we think that the notion of being *en rapport* is best thought of as displaying a kind of practical knowledge, a knowing who. For in that case singular propositions enter into the story in an inessential manner. It is interesting that at this point Kaplan reaches for the propositional attitude of knowledge. Notice how much less attractive is the thought "Are there two radically different ways of supposing a singular proposition, depending on whether you are *en rapport* with the object or not?" The semantic weight borne by 'know' is too great here. Being *en rapport* does not have any important theoretical role in delimiting the availability of singular thoughts.

We saw, when we accepted that we can have singular thoughts about Julius Caesar, that we can have singular thoughts even though we are not *en rapport* with the objects of those propositions. There are cases of being *en rapport* in which singular propositions are not involved. For example, I might use the expression

'The way things look to me'

as a description which picks out a certain way things strike me. Of course, we could introduce a name on the back of such a description, in the way Gareth Evans discusses.[25] That name would enable us to express thoughts not expressible before the name was brought into the language, unless we muddy the waters with modal operators and the like within the individuating description.[26] But the central point remains: individuating descriptions need not be more than actually individuating (and usually need not even be that) before we accept that the agent knows which object is in question. Certain sorts of practical ability which may be very context specific seem vitally important.

V. CONCLUDING REMARKS

I have distinguished between singular thoughts, that is, thoughts with a content whose truth requires that a particular object have certain properties, from *de re* belief, beliefs which are of an object brought before the mind and unmediated by perspective. I have argued that we have singular thoughts even while there are no *de re* beliefs.[27] Since it is always possible to have two thoughts about an object without knowing that it is the one object of those two thoughts, our beliefs are never belief *de re* properly so called, but since they are beliefs about particular objects they are singular thoughts. Let me reiterate and re-emphasize that though there are no *de re* beliefs, we do often legitimately use *de re* attributions of belief. These attributions can be apt and true without their marking the presence of a peculiar sort of belief. These are particularly useful ways of attributing a belief to someone, since they enable us to give an indication of the content

[25] Evans, "Reference and Contingency," *Monist*, LXII, 2 (1979): 161–89; reprinted in his *Collected Papers* (New York: Oxford, 1985), pp. 178–213. A word of warning: Evans's insistence that the epistemic status of an interpreted sentence be invariant across the language was introduced as seemingly innocent simplification. Our discussions of ## show that it was mistaken. Indeed it seems this is at the root of the mistaken supposed analysis of the notion of the *a priori* as the necessity of the diagonal in two-dimensional semantics. See my "The Problems with Double-Indexing Accounts of the *A Priori*," *Philosophical Studies*, CXVIII, 1/2 (March 2004): 67–81.

[26] But note that the description in question has none of these devices.

[27] I should emphasize that the denial of beliefs *de re* speaks not at all to the issue of *de re* necessity. In fact I hold that *de re* necessity makes very good sense.

of a person's belief while abstracting away from some aspects of the content. Again, *de re* attribution of belief is quite independent of the fact that we often have singular thoughts. So having singular thoughts is quite separable from having a belief attributed to you *de re*. Beliefs can be attributed *de re* without the underlying belief itself being singular, and the attribution can be made using a sentence which is not singular while the belief itself is singular. There is simply no requirement that belief attribution march in step with the character of the beliefs themselves.[28] Moreover, we are often truly said to know who someone is, but this is no particularly special ability. The Masked Man paradox shows us that to know who someone is cannot require that we recognize him in all contexts. In many contexts, an exam for example, providing a name will suffice. We do naturally contrast that sort of ability with being acquainted with or knowing the person, but even when we know a person, we need not have the ability to distinguish that person from all possible others, or even all actual others. The grounds for attribution of knowing who or which to someone are practical, not semantic. For that reason we should not expect to be able to find a kind of term whose occurrence within a belief attribution licenses exportation. Exportation is a third-person phenomenon and depends on the third person's views about the subject's beliefs.

We started this discussion with Hart's criticisms of Plato's attempts to characterize knowing as a grasping of being. The emphasis Plato gives to the firmness of the grasp, its unshakability, only serves to remind us of the anxiety that this account of knowledge is trying to displace. We can know that we know since our grasp on reality is so firm. The account of knowledge by acquaintance offered by Bertrand Russell responds to a similar anxiety, one inspired by the threat of a radical skepticism. The idea that the very possibility of having a thought requires knowledge by acquaintance, a form of knowledge which guarantees the existence of its purported objects, is suggested as an antidote to the skeptical worries. However, if I am right, such an antidote is a fraud. If we have a singular thought about an object, then that object exists, but we have no guarantees that what we take to be singular thoughts indeed are singular.

MICHAELIS MICHAEL

The University of New South Wales

[28] This point is independent of the denial of the existence of beliefs *de re*. Once we take off the theoretical blinkers we see that such attributions are actually very widely used, and used in a manner which is explicable in a satisfying manner. Even someone who holds that there nevertheless are beliefs *de re* must then suggest that only some *de re* attributions relate to beliefs *de re*.

A TOPOLOGICAL SORITES*

The paradigmatic cases of the sorites paradox—heaps of sand and bald heads—are cases where the changes in question are small but discrete. Trading on the vagueness of 'heap' and 'bald', we remove discrete units, grain-by-grain and hair-by-hair, to produce the paradox. Moreover, in the case both of the heap and of baldness, there is a natural ordering, in terms of the number of grains of sand and the number of hairs. Let us call such versions of the sorites paradox *discrete* and *numerical*. Most of the discussion in the literature concerns such cases, and why not? They are very difficult to solve and have lead to extremely fruitful work in philosophy of logic. But it is important to bear in mind that these are not the only versions of the sorites.

For a start, there are continuous versions of the sorites. Consider a sorites argument using the predicate *tall* and starting with someone who is 200 cm high and progressing *continuously* down to someone 125 cm high. Such versions of the paradox are usually shoe-horned into the above discrete format by considering a particular series of discrete transitions—1 mm steps, for example. Be that as it may, the underlying space here is continuous (putting aside, for the purposes of argument, debates about the discreteness or continuity of space-time). It seems that we ought to be able to formulate the sorites paradox in terms of continuous transitions and not merely discretize continuous cases. Indeed, it would seem that the smaller the increments, the more compelling the sorites argument, so the continuous version might well be thought to be the most compelling of all; see section II.

Next consider non-numerical cases of the sorites.[1] Here we have familiar examples of family resemblance concepts such as religion and sports. Consider an example of transitions from Hinduism (with its ritualistic dress and behavior, belief in supernatural beings with special powers, the passion-play of good versus evil, and a catalogue

* We are indebted to James Chase, Lloyd Humberstone, Dominic Hyde, and audiences at the Melbourne Logic Group and the AAPNZ 2009 for useful conversations on the topic of this paper. Research for this paper was funded by an Australian Research Council Discovery Grant to Mark Colyvan and Dominic Hyde (grant number DP0666020).

[1] Otávio Bueno and Mark Colyvan, "Just What Is Vagueness?" *Ratio*, XXV (2012), forthcoming.

0022-362X/10/0706/311–325

of hymns and chants); through the passionate following of an Australian Rules Football team (with slightly less ritualistic dress and behavior, belief in players blessed with extraordinary, if not superhuman, powers, the various heroes and villains, and the various chants and team songs); to a casual children's game of ball in a backyard. It is plausible that a sorites argument can be constructed here, but there is no natural ordering as in the numerical versions described at the start.

To take a more scientifically significant example, consider the concept of *endangered species*. Here we construct a transition from an abundance of some species, with ample connected habitat and with no population decline, through to a single individual member of some species, with little or only fragmented habitat and suffering rapid decline.[2] Both vagueness and its sorites paradox seem to be active in such a case, even if there is no salient textbook construction of a sorites argument ready to hand.

It is sometimes said that such family-resemblance cases are cases of higher-dimensional sorites, whereby each dimension (for example, the degree of ritualistic behavior) is well ordered, but there is no overall total ordering of the transition states.[3] Even this much numerical ordering strikes us as implausible, but be that as it may, cases such as this do not naturally lend themselves to the standard presentations of the sorites, and all the more so for other apparently vague notions, such as jokes, wisdom, or love. At the very least, we need to do violence to the case in order to get it to fit the standard discrete, numerical schema.

It is easy to set aside such cases or to insist that they conform to the discrete, numerical schema via suitable adjustments. In this paper, at least, we are not denying that such moves can be made. We are, however, questioning the wisdom of such moves. After all, on the face of it we have several quite different versions of the sorites. It may be that a narrow focus on the discrete, numerical versions such as the heap of sand obscures what really drives the paradox. Such a narrow focus may even lead to overconfidence in a solution that deals only with

[2] Helen M. Regan, Mark Colyvan, and Mark A. Burgman, "A Proposal for Fuzzy International Union for the Conservation of Nature (IUCN) Categories and Criteria," *Biological Conservation*, XCII (2000): 101–08; Regan, Colyvan, and Burgman, "A Taxonomy and Treatment of Uncertainty for Ecology and Conservation Biology," *Ecological Applications*, XII (2002): 618–28.

[3] Arthur W. Burks, "Empiricism and Vagueness," this JOURNAL, XLIII, 18 (Aug. 29, 1946): 477–86; Dominic Hyde, *Vagueness, Logic, and Ontology* (Burlington, VT: Ashgate, 2008), p. 17.

the special cases under consideration. It is at least plausible that the underlying phenomenon has little to do with discreteness or numerical ordering. Clearly, we would like a unified solution to the sorites; in order to achieve this, we first need a characterization of the sorites paradox in its full generality. Only then can we be confident that we are in a position to see what makes it tick.[4]

In this paper, we propose to provide such a general characterization. We will start with the canonical presentation of the sorites, then outline a continuous version, and then move to an even more general topological version of the sorites. The topological formulation is interesting in its own right, but it also leads very naturally to a new, more general definition of the problematic concept of *vagueness.*

The logic is classical throughout the paper, and the theorems are text-book. Accordingly, $\vdash$ represents classical consequence, and $\supset$ is the material conditional. All the proofs are well known and are given in thumbnail or omitted altogether. The message is that, just as classical logic and number theory make unbelievable predictions in canonical forms of the sorites, so classical topology makes exactly the same kind of paradoxical predictions in the more general case.

I. DISCRETE SORITES

Hyde offers a useful classification of the sorites paradoxes.[5] The first and most familiar is a long series of (material) conditional statements, with a true first sentence (0 grains is not a heap), seemingly true subsequent sentences (either 250 grains is a heap, or 251 is not), and a false conclusion (10,000 grains is not a heap). The second form of the sorites paradox is a generalization, called the *inductive* form. Let Φ be a predicate and $n \in \mathbb{N}$.

$$\begin{array}{ll} & \Phi 0, \\ & \forall n(\Phi n \supset \Phi(n+1)). \\ \vdash & \forall n \Phi n. \end{array}$$

This is just the mathematical induction schema. When Φ is a vague predicate the premises seem true, and this leads to trouble because a vague predicate is tolerant to small changes but does not apply to every object.

[4] Colyvan, "Vagueness and Truth," in Heather Dyke, ed., *From Truth to Reality: New Essays in Logic and Metaphysics* (New York: Routledge, 2008), pp. 29–40.

[5] Hyde, "Sorites Paradox," *Stanford Encyclopedia of Philosophy*, ed. Edward Zalta (2008). URL: http://plato.stanford.edu/entries/sorites-paradox/.

Since in the case of vague Φ the conclusion is false, we must reject the induction step (also called the sorites premise), and we thus arrive at the *line-drawing* form:

$$\begin{array}{ll} & \Phi 0, \\ & \neg\forall n \Phi n. \\ \vdash & \exists n(\Phi n \wedge \neg\Phi(n+1)). \end{array}$$

This is a valid argument with true premises, but it is still taken to be a paradox because it seems implicit in the notion of vagueness that a vague predicate cannot be sensitive to very small changes. And yet the line-drawing form concludes that there is a single second, a single grain of sand or hair on the head, that leads from being Φ to not; some straw breaks the camel's back.

In what follows, we will look for arguments analogous to these inductive and line-drawing forms which do not trade on the discrete ordering of $\mathbb{N}$. Let us emphasize that we are not challenging the legitimacy of the canonical sorites paradox qua paradox. Rather, we are looking for more abstract renderings that reveal the canonical sorites to be special cases of a more sweeping phenomenon.

II. CONTINUOUS SORITES

James Chase has generalized the sorites to the continuous case. His argument draws out consequences of the distinctive axiom for continuity, which is as follows.[6]

> *Axiom 1 (Dedekind)* Let $A \cup B = \mathbb{R}$ be nonempty and disjoint sets, with $a < b$ for every $a \in A$ and $b \in B$. There is a unique $k \in \mathbb{R}$ such that $a \leq k \leq b$ for every $a \in A$ and $b \in B$.

From Dedekind's axiom we have the (equivalent) proposition that any set of reals bounded from above has a least upper bound. Then, by a standard series of lemmas, beginning with the Archimedean property (that for all $x \in \mathbb{R}$ there exists an $n \in \mathbb{N}$ such that $x < n$), it follows that the reals are dense, in the sense that if $x < y$ then there is a real z such that $x < z < y$. So much for how Dedekind's axiom constitutes the reals.

Consider a vague predicate Φ mapped onto a real-number interval [0, 1], exhaustively partitioned into two nonempty sets,

$$\begin{aligned} A &= \{x \in [0,1] : \Phi(x)\}, \\ B &= \{x \in [0,1] : \neg\Phi(x)\}, \end{aligned}$$

[6] We are already assuming the other usual definitions and field properties of the real numbers $\mathbb{R}$, as can be studied in any text, for example, Michael Spivak, *Calculus*, 3rd ed. (New York: Cambridge, 2006).

with $a < b$ for all $a \in A$, $b \in B$. We assume that $\Phi(0)$ and $\neg\Phi(1)$, and that if some number is not Φ, then no numbers after it are Φ either. Thus A is the left side of the interval and B is the right. The left set has a least upper bound; call it *supA*. Now, Φ is vague, and in discrete cases we are prepared to admit that objects differing by whole number amounts (a hair, a grain of sand) are too similar for one to be Φ but not the other. Here the objects in question are much closer together. Therefore, since points vanishingly close to *supA* are Φ, and Φ is vague, also $\Phi(supA)$. By a symmetrical argument, $\neg\Phi(inf\,B)$. Knowing this we have a paradox.

By the linear order on $\mathbb{R}$, one of the following must be true:

$$\begin{array}{rl} & supA < inf\,B \\ \text{or} & inf\,B < supA \\ \text{or} & supA = inf\,B. \end{array}$$

Since the reals are dense, we have the following contradiction. If *supA* and *inf B* are different numbers, then there is some z between them, $supA < z < inf\,B$ or $inf\,B < z < supA$. But then Φz and $\neg\Phi z$, since by definition anything less than *inf B* is Φ but anything greater than *supA* is not. On the other hand, if $supA = inf\,B$ then again $\Phi supA$ and $\neg\Phi supA$. This exhausts all the cases. Therefore there is a point both Φ and $\neg\Phi$, a contradiction.

The Dedekind axiom can be used to derive the intermediate value theorem, and here we just have a special case of this. A continuous path must cross over from A to B at some distinct point. The transition is problematic if the sets are supposed to be partitioned by a vague property.

The argument used by Chase can be represented in analogy to the discrete inductive form. A sequence $\mathcal{X} = \{x_0, x_1, ...\}$ is *Cauchy* iff for all real ε there is some $n \in \mathbb{N}$ such that $|\, x_i - x_j \,| < \varepsilon$ as long as i, $j > n$. Let $\mathcal{X}$ range over Cauchy sequences in the interval [0, 1]. The soritical argument now runs:

$$\begin{array}{rl} & \Phi 0, \\ & \forall \mathcal{X}(\forall x(x \in \mathcal{X} \supset \Phi x) \supset \Phi(sup\mathcal{X})), \\ \vdash & \Phi 1. \end{array}$$

The second premise is the sorites premise. This is not entirely analogous to the discrete case, since this is not a generally valid mathematical schema. Priest calls it the Leibniz continuity condition: whatever is going on arbitrarily close to some limiting point is also going on at the limiting point; *natura non facit saltus.* Were it generally valid, we could prove all sorts of nonsense.[7] In the case of a vague predicate, though, the condition seems ineluctable. Since it causes trouble, similarly to the discrete case, we negate the sorites premise and get a line-drawing form:

$$\begin{array}{ll} & \Phi 0, \\ & \neg\Phi 1 \\ \vdash & \exists\mathcal{X}(\forall x(x \in \mathcal{X} \supset \Phi x) \wedge \neg\Phi(sup\mathcal{X})), \end{array}$$

again where $\mathcal{X}$ is Cauchy.

What can we learn from this version of the paradox? For a start, we see how the sorites can be constructed so that it relies upon a property of the real line—the property of being *connected*. This property can be expressed with the notion of *metric adherence* (where topological adherence is defined in section IV below): A point x is *adherent* to a set X iff for any ε, no matter how small, the ε-sized interval around x includes points in X. With this in hand, we see, just as a consequence of Dedekind's axiom, that the interval [0, 1], and the reals in general, cannot be broken into two isolated parts:

Theorem 1 If $\mathbb{R}$ is partitioned into two nonempty, disjoint sets, some number is adherent to both sets.

Proof. Let $A \cup B = \mathbb{R}$, with $a \in A$ and $b \in B$. Without loss of generality suppose $a < b$. Then $inf\{x \in B : a < x\}$ is adherent to both A and B.□

A very common response to the discrete forms of the sorites paradox is to see a problem with exclusively and exhaustively separating objects into two categories, Φ and not. We now see that this problem is well expressed in terms of connecteness. Connectedness as exemplified in Theorem 1 is an emergent property of Dedekind's axiom, and the key in generalizing from the discrete to the continuous. We can now use this property to generalize again.

III. A TOPOLOGICAL SORITES

For millennia, geometers attempted to prove Euclid's parallel postulate. In the late eighteenth century came awareness that there are models of the first four Euclidean axioms that do not respect the parallel postulate. By the nineteenth century, in his landmark paper on the foundations of geometry, Riemann was able to diagnose *why* there are such models: The first four postulates, he saw, codify topological properties of the space, while the fifth is a specifically metric

[7] Graham Priest, *In Contradiction: A Study of the Transconsistent* (New York: Oxford, 2006), chapter 11.

property.[8] The lesson from Euclid is that there is a distinct science of space that does not deal in metric, quantitative notions, but only in qualitative notions like closeness.

It will be useful to describe the standard concepts of point-set topology.[9] The basic primitive (though intuitively familiar) notion is that of *open set*. Let X be a set. A *topology* is a collection of open subsets of X, closed under union and finite intersection, and including X and the empty set $\emptyset$. Let A be a member of the topology on X. A point x is interior to A, and A is a *neighborhood*[10] of x, iff there is an open set U where $x \in U \subseteq A$. A set A is open iff all its points are interior, that is, A is a neighborhood of all $x \in A$.

The interior of A is its largest open subset, the union of its open subsets, A°. The closure of A is its smallest closed superset, the intersection of closed supersets, A^-. The interior, the set, and the closure sit like this:

$$A^\circ \subseteq A \subseteq A^-.$$

A set A is open if A is contained in its interior, $A \subseteq A^\circ$, and A is closed if A contains its closure, $A^- \subseteq A$. Therefore a set is *both open and closed* if $A^\circ = A^-$.

Definition 1 A space X is *connected* iff the only sets in the topology of X that are both open and closed are X and $\emptyset$.

The following consequence could serve equally well as the definition of connectedness.

Theorem 2 A space is connected iff it cannot be partitioned into non-empty, disjoint, open sets.

At Theorem 1, for example, we saw that the reals $\mathbb{R}$ are connected.[11] We are now in a position to say why connected spaces are so useful for our present purposes.

[8] This and other insights are explored in Michael Spivak's *A Comprehensive Introduction to Differential Geometry*, vol. 2 (Berkeley: Publish or Perish, Inc., 1979), chapter 4.

[9] A standard reference is John L. Kelley's *General Topology* (New York: Springer-Verlag, 1955).

[10] The notion of a neighborhood is due to Hausdorff. He used the word *die Umgebung*, hence the common use of the symbol 'U'.

[11] There is a stronger notion, of a *path-connected space*, in which every two points a, $b \in A$ are connected by a path, a continuous function f: $[0, 1] \rightarrow A$ with $f(0) = a$ and $f(1) = b$. Every path-connected space is connected, but a connected space can still be impassible between two points (for example, the "topologist's sine wave"). See Lynn Arthur Steen and J. Arthur Seebach, Jr., *Counterexamples in Topology* (New York: Springer-Verlag, 1978). In multi-dimensional cases of vagueness, path connectedness seems to be the property that generates the paradox: We follow an arbitrary path through the space which takes us monotonically from one point in the space another. In pursuit of full generality, however, we will stick with the more general notion of connectedness.

Definition 2 A function f is *locally constant* iff for each $x \in X$ there is a neighborhood U_x such that the restriction of f to U_x is constant. A *globally constant* function always takes the same value, without restriction.

This is the key lemma.

Lemma 1 Let X be a connected space, Y a set, and f a function from X to Y. Suppose that f is locally constant. Then f is globally constant. A fortiori, if y is in the range of f, then $X = \{x : f(x) = y\}$.

Proof. Suppose f is not globally constant. Then there are objects x, $y \in X$ such that $f(x) \neq f(y)$. Then there is a $z \in X$ such that for any of its neighborhoods U_z, there are objects $x, y \in U_z$ and $f(x) \neq f(y)$.□

Heuristically, the set Y in Lemma 1 can be thought of as the pair $\{0, 1\}$, in which case the *characteristic function* σ of the set A is defined thus:

$$\sigma_A(x) = \begin{cases} 1 & \text{if} \quad x \in A, \\ 0 & \text{if} \quad x \notin A. \end{cases}$$

Consider a predicate Φ mapped onto a set A, and say that A is the extension of Φ. We have the analogous

$$\sigma_\Phi(x) = \begin{cases} 1 & \text{if} \quad \Phi(x), \\ 0 & \text{if} \quad \neg\Phi(x). \end{cases}$$

Since this set-up will lead to a paradox, there could be some objection to the language just employed: In using sets to represent predicates, we are assuming (with classical model theory) that predicates have extensions and these extensions are sets. We are eliding between predicates and sets, and it could be pointed out that, ever since Russell told Frege, we have known this is not always a harmless elision. There is, however, good reason to neglect such distinctions in this paper. The reason is that our goal is merely to formulate a problem using what looks like, in other cases, unproblematic language—to state, without jumping to solve, a paradox. We help ourselves to talk about extensions, only flagging that this language is not entirely innocent; see section v.

We use the notions of local constancy and characteristic function to propose a definition of vagueness.[12]

Definition 3 (*Vagueness*) A predicate is *vague* iff its characteristic function is locally constant but not globally constant.

The definition says that a vague predicate is tolerant of small changes but does run out somewhere. The principle of tolerance is found in the

[12] Thanks to Lloyd Humberstone for his contribution to formulating this definition.

standard literature on the sorites.[13] All the same, this definition is quite unlike any of the usual definitions in the literature.[14] But this is to be expected for two reasons.

First, it is well recognized that it is extremely difficult to provide a definition of vagueness that does not beg questions about its proper treatment.[15] For example, a common definition of vagueness in terms of permitting borderline cases, which in turn are defined as gaps, begs the question against glutty approaches.[16] While it is not the purpose of the present discussion to defend the above definition of vagueness against all charges of being question begging, its generality does suggest that it will do better on this front than some of the others—at least it does not presuppose that vagueness is a gappy rather than a glutty, or even nonclassical, phenomenon.[17] In any case, it is a very natural definition in the context of a more general conception of vagueness and is worth laying on the table.

This brings us to the second reason it is not surprising that this new definition is different from the standard ones: the standard definitions have a much narrower phenomenon as their targets—typically, vagueness associated with discrete, numerical sorites. Our aim is to provide a more general account of the sorites, and this must be accompanied with a more general definition of vagueness. Often, generalizations lead to new and more fecund definitions of the target concepts.[18] Still, we need to show that this definition does capture the intuitive notion. We do this

[13] Crispin Wright, "On the Coherence of Vague Predicates," *Synthese*, XXX, 3/4 (April–May 1975): 325–65.

[14] See for examples: Kit Fine, "Vagueness, Truth and Logic," *Synthese*, XXX, 3/4 (April–May 1975): 265–300; Hyde, "Sorites Paradox," *op. cit.*; Rosanna Keefe, *Theories of Vagueness* (New York: Cambridge, 2000); Roy Sorensen, *Vagueness and Contradiction* (New York: Oxford, 2001); Stewart Shapiro, *Vagueness in Context* (New York: Oxford, 2006); Nicholas J. J. Smith, "Vagueness as Closeness," *Australasian Journal of Philosophy*, LXXXIII (2005): 157–83; Timothy Williamson, *Vagueness* (New York: Routledge, 1994); Crispin Wright, "On the Characterisation of Borderline Cases," forthcoming.

[15] Shapiro, *op. cit.*; Bueno and Colyvan, *op. cit.*

[16] Hyde and Colyvan, "Paraconsistent Vagueness: Why Not?" *Australasian Journal of Logic*, VI (2008): 107–21; Zach Weber, "A Paraconsistent Model of Vagueness," *Mind*, to appear.

[17] Substituting 'continuous' for 'constant' in the definition, which would make no great difference in what follows, a fuzzy account can also be allowed for. Smith briefly entertains a definition of vagueness that does just this: A predicate is vague if its characteristic function is continuous (*Vagueness and Degrees of Truth* (New York: Oxford, 2008), p. 182). Smith works with degrees of truth (cf. his "A Plea for Things That Are Not Quite All There: Or, Is There a Problem about Vague Composition and Vague Existence?" this JOURNAL, CII, 8 (August 2005): 381–421); to make his arguments relevant to our restriction to truth values of only 0 or 1, we would say the characteristic function is constant. Smith points out some difficulties with a topological theory of vagueness, and this proposed definition in particular (*ibid.*). Since his objections are tied to continuity per se they do not impact the proposal here. We briefly take this up in section V.

[18] For example, to generalize the concept of a *straight line* to that of a *geodesic*, one does not use the seemingly obvious idea of the shortest distance between two points,

by way of a couple of formal results and show that it can be used to formulate different versions of topological sorites arguments.

Theorem 3 Let X be connected and $A \subseteq X$, with A the extension of a vague Φ. Then either $A = X$ or $A = \emptyset$.

Proof. The characteristic function σ_A of a vague predicate A is locally constant, by Definition 3. So, by Lemma 1, $\sigma_A : X \longrightarrow \{0, 1\}$ is globally constant.□

What does this mean? In general, for a connected space X, to prove that every $x \in X$ has some property Φ it suffices to show

BASE: Some $x \in X$ is Φ, and
INDUCTION: x is Φ iff all the points in a sufficiently small neighborhood of x are Φ.

These establish that all $x \in X$ are Φ because the induction asserts that the characteristic function for Φ is locally constant, and the base step asserts that the global value is not 0. This gives us a *topological inductive* version of the sorites:

Sorites Paradox, topological inductive version:
Let Φ be a vague predicate and X be a connected space.
Then if any member of X is Φ, every member of X is Φ.

Alternatively, with A as the extension of vague Φ,

$$\begin{aligned} & \emptyset \neq A \subseteq X, \\ & X \text{ is connected.} \\ \vdash \quad & A = X. \end{aligned}$$

The 'induction step' is that X is connected, because connected spaces support the local-global property of Lemma 1, now built into the definition of vagueness. Faced with the discrete inductive sorites, we reject the induction step. Similarly, here we reexamine the situation from a line-drawing perspective.

Theorem 4 For any X such that $A \subset X$, that is, any space containing A other than A itself, if A represents a vague predicate then X is disconnected.

Proof. The characteristic function on A is locally constant, by Definition 3. Therefore it is globally constant, by Lemma 1. Now, $A = \{x : \sigma_A(x) = 1\}$. If X is connected, then $X = \{x: \sigma_A(x) = 1\}$, too; so if $X \neq A$, then X must be disconnected.□

---

but rather that of a curve whose acceleration is identically zero. The latter turns out to be more flexible and informative. See John M. Lee, *Riemannian Manifolds: An Introduction to Curvature* (New York: Springer, 1997), p. 47.

The result is a *topological line-drawing* sorites:

Sorites Paradox, topological line-drawing version:
Some things in X are Φ and some are not, for some vague Φ.
Then X is disconnected.

Alternatively, again with A as the extension of Φ,

$$\begin{array}{ll} & \emptyset \neq A \subseteq X, \\ & A \neq X. \\ \vdash & X \text{ is not connected.} \end{array}$$

Recall that, for the moment, we are only finding natural formulations of the paradox, in analogy to the discrete cases, without insinuating anything about how to interpret or resolve the problem. (It is interesting to notice how difficult it is only to state a problem, without trying to solve it.) The aim of the exercise is to show how topology represents vagueness, under classical assumptions—for example, that predicates can be represented with extensions. Under this assumption, classical topology predicts that the host-space of a vague predicate is not connected, because otherwise vague predicates would apply to everything.

IV. A CLOSER LOOK AT BOUNDARIES

Connectedness is a global property; it cannot be determined locally. But disconnectedness is a very local property. If σ is not globally constant, then it is not locally constant, either (by contraposition). So the disconnection becomes a local property, and the familiar counterintuitive aspect of line-drawing emerges. There is an $x \in A$, some particular point, in the neighborhood of which σ_A changes value.

Owing to the extreme locality of disconnection, we can study vagueness by studying the behavior of the characteristic function at the boundary of A. What would the boundary have to be like to support a sorites? Let us have a closer look at the boundary of a space.

In the following, the complement of A is $\mathcal{C}A = \{x : x \notin A\}$ and $X - A = X \cap \mathcal{C}A$.

Definition 4 A point $x \in X$ is adherent to A iff every neighborhood of x in X intersects A. The *boundary* of A, $\partial(A) := A^{-} \cap (X - A)^{-}$, is the set of all points adherent to both A and $X - A$.

A set shares its boundary with its complement, $\partial(A) = \partial(X - A)$, and a boundary is 'stable' in the sense that $\partial(\partial(A)) \subseteq \partial(A)$. Moreover, unpacking definitions,

$$\begin{aligned} \partial(A) &= A^{-} - A^{\circ}, \\ A^{\circ} &= A - \partial(A), \\ A^{-} &= \partial(A) \cup A^{\circ}. \end{aligned}$$

If $A \subseteq X$, then X is the pairwise disjoint union of $\partial(A)$, A^{-}, and $\mathcal{C}A^{-}$.

Now we have an alternative characterization of fundamental notions, summarized in a theorem:

Theorem 5 A is open iff $\partial(A) \subseteq X - A$, and A is closed iff $\partial(A) \subseteq A$.

A closed set includes all its adherent points; the closure of A is the set of all points adherent to A; so A is closed iff every point adherent to A is in A. It should now be clear that *set-theoretic extensions of vague predicates* are closed. For example, in the continuous case in section II, both the left set and the right set were closed by virtue of vagueness. Our further assumption that the sets be disjoint induced a contradiction.

We have now collected enough tools to make the last more precise, and to draw some morals. From the definition of adherence, it follows that a point k is adherent to A iff k is not interior to $X - A$, and k is interior to A iff k is not adherent to $X - A$. A fortiori, $\partial(A)$ is all the points interior to neither A nor $X - A$. In this terminology, A is open iff $A \cap \partial A$ is empty. Finally, A is both open and closed if $\partial(A) \subseteq A \cap (X - A)$. In sum,

Theorem 6 The following are equivalent:
The characteristic function on A is constant;
A is both open and closed, $A^\circ = A^- = A$;
$\partial(A) = \emptyset$.

This tells us that $x \in \partial(A)$ implies $x \in A$ and $x \notin A$, just as we saw in section II. Classically, this means that the boundary of A is empty, on pain of contradiction. From the definition of connectedness (Def. 1), the only sets with empty boundaries in a connected space X are X itself and $\emptyset$. Similarly, a connected space X is both open and closed, $X^\circ = X = X^-$; so if X is the set-theoretic extension of a vague predicate, then by the last theorem its boundary is overloaded. In section II, Chase derived a contradiction owing to the simple fact that if X is connected, $A \subseteq X$, and $B = X - A$, then A is open iff B is closed. The sorites premise made both sides of the interval both open and closed, overloading the boundary.

The sorites is a paradox. Classical theory encounters difficulties in the face of paradoxes, and such is the case here. The possibilities for a nonclassical treatment may be more flexible—for example, we might employ characteristic *relations* rather than functions, as is done in relational semantics.[19] On this approach, we could also organize things like $\vdash$ so that a glutty boundary is not disastrous. But to a fair extent these strategies await the development of more nonclassical mathematics. In the meantime, we have reached our destination, and step back to examine it.

[19] Priest, *An Introduction to Non-Classical Logic* (New York: Cambridge, 2008).

V. OPEN QUESTIONS

A number of questions arise. For now we consider an interconnected few concerning extensions and the representation of vagueness.

Let Φ be our vague predicate, and let A be the set-theoretic extension of Φ. One might wonder, following Smith,[20] as to whether there is a topology on the domain of the characteristic function of Φ. Simply stipulating that there is a topology on the domain, he argues, may be an "onerous" assumption not grounded in empirical experience. It is true that generating a topology is not always successful. For example, working with his own definition of vagueness in terms of a three-place 'closeness' relation, Smith is only able to generate the *discrete* topology,[21] which trivializes the exercise. And it is true that it is not always obvious that the space of a predicate like 'is an endangered species' has a natural topological structure. On the other hand, it does not seem to us very presumptuous to assume a nontrivial topology on the domain of the characteristic function of Φ. This depends, to an extent, on what the domain is. But whether or not we provide a recipe for generating a topology, it is entirely plausible that there is one in many interesting cases. Nontrivial topologies are not so hard to come by. In many cases, topologies can be inherited from Euclidean space. Or, returning to the definitions of boundary, for example, $\{\mathcal{C}(A \cup \partial(A)) : A \subseteq X\}$ is a topology.

More fundamentally, we can ask whether or not the extension of a predicate can always be interpreted with sets. If we are taking extensions to represent predicates, then this is asking whether or not vagueness can be represented by standard model theory. Perhaps, as has been the solution to Frege's naive comprehension woes, the answer here is to deny that every extension is a set. Perhaps extensions of vague predicates are not sets. However, this is a most unappealing thought, once we notice the large number of vague predicates in both our conversational language and more rigorous scientific language. The predicates in question here are not unusual like 'is not a member of itself' or 'is an ordinal', as in the Russell and Burali-Forti paradoxes, but banal predicates about sports and loud noises. We have been assuming since Descartes that the world is uniquely quantifiable by maps from the world to the real numbers; we have been assuming since Einstein that differential geometry and tensor calculus provide the tools to understand macroscopic space and time. These assumptions rest on basic representations of the world via set theory. If it turns out that vague properties cannot be treated in this way, then that would be very, very surprising.

[20] Smith, *Vagueness and Degrees of Truth*, p. 152.
[21] See Kelley, *op. cit.*, p. 37.

With set-theoretic extensions looked at in this way, then, there is a further open question. Is the topological characterization of sorites paradoxical at all? Candidates for multi-dimensional vagueness are commonplace but are more abstract than predicates like redness or baldness. The space of jokes, wisdom, or love may well be disconnected, being abstract fragments of logical space to begin with. Was there reason to have suspected otherwise? Rather than a paradox, perhaps in these cases we have learned something structural—that, in a sense, the transition from religion to sport involves traversing a disconnection.

Nevertheless, it is also clear that there will be at least some cases where multi-dimensional sorites traverse abstract connected spaces. Some of these spaces (for example, Minkowski space) will be of scientific interest, and there will be sorites arguments in these spaces which will be hard to represent as examples of canonical discrete, numerical sorites. The space of threatened species, for example, has several degrees of freedom: species numbers, area and quality of habitat, rate of decline, and others. While each of these is arguably numerical, it is not at all clear how to combine them in a meaningful fashion. Any proposed metric in this space, it seems, will be problematic. Yet the space in question is very plausibly connected. Importantly, whether or not this and other spaces are unexpectedly disconnected, we have made progress, because the characterization we have given here will help in formalizing the alleged sorites arguments in all such cases. Our topological characterization leaves open the question of whether the space really is connected; it leaves the question sharpened.

Perhaps the concern can be put differently—as a concern about maintaining a neutral dialectic. We have proven that whenever we have a sorites series, the underlying space must be disconnected. This should be music to the ears of epistemicists about vagueness, for they take the lesson learned from the sorites paradox to be that all such spaces are (surprisingly) disconnected. It seems we have just vindicated the epistemic approach and thus trivialized the debate by ruling out other serious contenders such as supervaluational and fuzzy approaches. Any definition that begs all the important questions is no definition at all.

This, however, is to misunderstand where things stand. As we just indicated, some of the spaces in question are connected—some are provably connected (for example, Minkowski space and $\mathbb{R}^n$), while others have strong cases to be made for their connectedness (for example, the space of threatened species). This is the heart of the sorites and why it is a paradox. Just as vague predicates in the company of classical logic lead to genuine paradox, so too does classical topology with vague set descriptions. All along, our purpose has been to provide a generalization of the sorites; we have not been trying to solve the paradox. Of course if one rigidly sticks with classical topology, then epistemicism appears to be the only viable approach; this is no different from the canonical sorites, where if one refuses to depart from classical logic, epistemicism appears to be the only viable option. But all the standard moves are here to be made; no proposed solution is ruled out. For example, a very natural move to make here is to entertain a nonclassical theory of topology, one underwritten by a nonclassical set theory—fuzzy, intuitionisitic, or paraconsistent—or through new directions in topological computation.[22] Our discussion of a generalized sorites via an appeal to classical point-set topology was not intended to suggest that classical topology should be used in solving the problem.

We do not expect that we have said enough here to convince everyone that the sorites is essentially topological. An important concern is whether the generalization is too general. A topological characterization of vagueness is a generalization on the canonical discrete, numerical sorites in the same way that topology itself is a generalization on certain properties of the real numbers. One important consequence of the topological approach is that, since a topological sorites does not require assumptions about order, any proposed solutions to sorites that trade on the details of order are not going to be general solutions. Conversely, this could mean that the generalization has gone too far and has lost grip on the essence of the sorites. If so, topology is the wrong tool for the job, and what we are calling a sorites is, in fact, not.

If, on the other hand, the argument form we display with only topological tools is still a recognizable sorites, then the 'lost' information is inessential. With some doubts now aired, we do think that a topological sorites is recognizably a generalization of the canonical sorites and that the topological characterization captures the essential ingredients—namely, connectedness and local and global constancy. It is not hard to see that the core notion of local constancy is a generalization of the principle of tolerance, and that the topological sorites is a generalization in the sense that the canonical cases can be recovered as special cases. If we are right about this, then progress has been made in exposing what is invariant about vagueness in a variety of cases, and we are closer to understanding a very resilient puzzle.

ZACH WEBER

University of Sydney and University of Melbourne

MARK COLYVAN

University of Sydney

[22] Viggo Stoltenberg-Hansen and John V. Tucker, "Computability on Topological Spaces via Domain Representations," in *New Computational Paradigms: Changing Conceptions of What Is Computable* (New York: Springer, 2008.)

COMMENTS AND CRITICISM

ON REASONS AND EPISTEMIC RATIONALITY*

In a recent paper, John Hawthorne and Jason Stanley propose an analysis of how knowledge and action relate to each other.[1] According to their *Reason-Knowledge Principle* (RKP), it is appropriate to treat the proposition that p as a reason for acting iff we know that p, for p-dependent choices;[2] within their account, in addition, knowledge delivers probability 1. Hawthorne and Stanley also note that sometimes it is intuitively rational to act on partial beliefs. What is appropriate to treat as one's reason for action, in this case, is the epistemic probability of p conditional on the agent's total knowledge K. RKP then requires that one knows that $P\ (p \mid K) = r$ (for some r).[3]

Here I am not concerned with the general formulation of RKP, but I shall focus exclusively on the more restricted contention according to which agents should not invoke probability claims as reasons unless they know that such claims are true. I believe there are grounds to think that this is a problematic demand.

I would like to begin by reflecting on the role of personal probabilities at the time of justifying action. The authors make it clear that RKP is meant to refer to *objective* probability functions. Still, they do not say explicitly whether, in their view, subjective measures can ever act as motivating reasons in their own right. At any rate, the linguistic evidence does not seem to exclude this possibility in any obvious way, as I hope to show below. In particular, note that the fact that a given probability claim is best interpreted as being epistemic does not mean it is not subjective, where subjective measures may very well incorporate estimates about (physical) chances. Unfortunately, if probability judgments are—at least at times—taken to encode personal measures, then RKP falls short of what we need.

* A previous version of this commentary was read at a workshop held by the *Grupo de Acción Filosófica* (GAF) at the *Universidad de Buenos Aires* in April 2008. I am indebted to Jason Stanley for valuable feedback. I also want to thank the members of GAF for discussion and comments.

[1] Hawthorne and Stanley, "Knowledge and Action," this JOURNAL, CV, 10 (October 2008): 571–90.

[2] A choice between options $x_1 \ldots x_n$ is said to be p-dependent iff the most preferable of $x_1 \ldots x_n$ conditional on the proposition that p is not the same as the most preferable of $x_1 \ldots x_n$ conditional on the proposition that not-p.

[3] This is my notation, not theirs.

0022-362X/10/0706/326–330

Indeed, Hawthorne and Stanley concede that RKP does not mesh well with personal probabilities (584). But the actual explanation as to why this is so is not addressed by their paper. The crucial point is that positing knowledge—or even belief—adds an unnecessary complication, and ultimately distorts the nature of the underlying phenomenon. In this respect, there seems to be an interesting analogy between credences and desires, preferences, possibility judgments, or aesthetic judgments. Suppose that, as far as *S* is concerned,

x is desirable;
it is correct to do *y*;
it would be nice if *p* obtained;
p is preferable to *q*;
p is possible;
p is highly probable.

In each case we can identify a primary attitude that consists in desiring a particular object, or preferring the occurrence of a particular state of affairs, or conceiving of the occurrence of a particular state of affairs as more or less probable. All such attitudes play a vital role in the economy of an agent's epistemic life, at the time of engaging in both theoretical and practical reasoning. Suppose now that *S* believes (knows) that she is committed to a particular set of personal probability judgments (desires, judgments of taste, and so on). At least in typical scenarios, the corresponding second-order belief (knowledge) will not do any real work in *S*'s acting or reasoning in a certain way, over and above what is already achieved by the first-order level.[4] Consider, by way of illustration: "Why have you moved your arm?" "Because I wanted to reach the bottle and drink some water" (rather than: because I believed/knew that I wanted to reach the bottle and drink some water); "Why have you bought that paint?" "Because I like it" (rather than: because I believe/know that I like it); "Why are you carrying an umbrella with you?" "Because it seems likely [to me] that it will rain" (rather than: because I believe/know that it seems likely that it will rain). Thus, if we are to trust our ordinary use of the language, the reason I moved my arm was a primary desire, so to speak, and not a second-order belief, or a piece of second-order knowledge, about my having a particular desire.[5] Likewise, the reason I am carrying an umbrella is a primary

[4] By the expression 'second-order belief' I mean to refer to a belief about a first-order attitude that itself may, but need not, be a belief; *mutatis mutandis* for 'second-order knowledge'.

[5] Of course, this is not to deny that a complete account of my reasons for acting may well incorporate, in addition, an array of *first*-order (full) beliefs, as well as further desires. (Thanks to Alejandro Cassini for pressing this point.)

probabilistic commitment, and not a belief, or a piece of knowledge, about a particular probabilistic claim.

Notice, moreover, that first-order knowledge claims behave very much unlike putative second-order propositional attitudes on personal probability judgments. “I invested in the market because the profit was going to be high” and “I invested in the market because I knew the profit was going to be high” are often interchangeable; “I invested in the market because it seemed very likely to me that the profit was going to be high” and “I invested in the market because I knew it seemed very likely to me that the profit was going to be high” are not—the last assertion is just awkward. Its awkwardness tells us something important about how partial beliefs enter into the business of giving and asking for reasons.[6] Once we admit personal probabilities into the picture (and it is not clear what would prevent us from doing so), motivating reasons might turn out to be other than the content of (true, justified) beliefs: primary probabilistic commitments can also do the trick. Thus, it seems perfectly right to treat a particular probabilistic commitment *C* as a reason for acting, without thereby requiring knowledge of *C*—or belief therein, for that matter.

It might be objected here that having *C* requires our knowing that we have it, out of rationality considerations. But this point is irrelevant for the present discussion: regardless of the intrinsic value of second-order attitudes, the examples presented above show that our reasons for acting typically can be found in primary commitments—epistemic and otherwise. To put it differently: if an agent has *C* but does not know she has it (say, because of transparency failure), she is already at fault; her further treating *C* as a reason does not add any extra offense. In short, if probabilistic talk is interpreted along subjectivist lines, RKP can be violated without intuitively making the agent accountable as far as her treatment of reasons is concerned.[7]

[6] Of course, we could always conceive of particular scenarios in which focusing on second-order attitudes becomes acceptable (“Are you sure you don’t want to try the cake?” “Yes—I *know* I don’t like chocolate”). But this is beside the point—the fact remains that mentioning second-order attitudes on probabilistic commitments, desires, or preferences is usually idle, and it very often leads to infelicities.

[7] Incidentally, to say that RKP *can* be violated without making the agent accountable is not equivalent to saying that our failure to know that we have *C* should never prevent our treating *C* as a legitimate reason. We might still contend, for instance, that treating *C* as a reason should imply, at the very least, that it is *true* that we have *C* (say, if we understand ‘reasons’ in the same objective way Hawthorne and Stanley do). In any event, notice that the situation here is not analogous to the one the authors have in mind when they present their principle in terms of full beliefs. Unlike the case in which an agent falsely believes a given proposition *p* about the external world, the agent who misidentifies her probabilistic commitments can be charged with irrationality, rather than with a mere factual mistake. Hence, once again, there is room to argue that treating *C* as a reason does not add further irrationality.

Let me turn now to Hawthorne and Stanley's preferred interpretation of probabilities. A theory of objective functions of the sort required by the authors would need to tell us how to obtain objective confirmation measures between p and K, for any possible p and K. But we are not given any hints as to how such a confirmation theory could go; more importantly, we are not given any reassurance that such a theory is possible in the first place. In the absence of any details, we seem to be left with how much S takes K to confirm p—but this, of course, takes us back to the realm of personal measures.

In any case, the authors' idea is that, ultimately, by focusing on the evidence we can circumvent mentioning probabilities altogether: acting on knowledge of epistemic probabilities would be tantamount to acting on the propositions on which we conditionalize in order to define the particular epistemic measures we have (584–85). To know that $P(p \mid K) = r$ is just to have K. If this were true, we would indeed get rid of the problem of providing a suitable interpretation for probability talk in the natural language, at least vis-à-vis an analysis of the link between knowledge and reasons.

But this move will not do. Consider an agent who asserts,

(*) "The reason I treated patient A with drug d on this occasion was that, as far as I know, drug d cured some people in the past and killed others. Moreover, several other untreated patients with symptoms similar to those of A died a horrible death."

Is this an admissible way for the agent to justify her action? Hardly so; we just cannot see where the motivation for her behavior lies. The awkwardness of the agent's discourse reveals precisely that we cannot assume probabilities to be implicitly operating here; it also shows that there is no straightforward route that could take an agent from K to knowing a relevant set of probability claims. Examples like (*) can be easily multiplied; except perhaps for extremely simple cases, the richer probabilistic structure that typically superimposes on K can be crucial at the time of deciding what counts as an appropriate motivating reason.

Perhaps the idea is that we should just assume the existence of a prior objective probability distribution that provides the input to calculate the relevant conditional measures.[8] Even though the authors never go down this path explicitly, we could seek to interpret their

[8] For instance, Timothy Williamson has proposed an objective sort of Bayesianism, according to which we can identify a prior probability distribution that measures the intrinsic plausibility of hypotheses prior to investigation. See his *Knowledge and Its Limits* (New York: Oxford, 2000), p. 211.

position as loosely based on this thought. However, according to this interpretation Hawthorne and Stanley should say that $P(p \mid K) = r$ can be treated as a legitimate reason for action only if the agent has knowledge of the objective priors on the basis of which suitable confirmation measures are obtained. But, as example (*) shows, in typical cases we cannot assume references to a hypothetical prior distribution to be implicit in standard discourse. Hence, at the very least, for the explanation of an action to make sense the agent would need to mention the relevant priors. Moreover, the lack of linguistic evidence can be taken to favor skepticism on the very existence of objective priors of the type required; postulating such measures might still turn out to be productive on theoretical grounds, but we are then left with the urgent task of discussing the details of such a theory and its relation with reasons as normally given by speakers—*precisely* because Hawthorne and Stanley place reasons at the center stage of their project.

To sum up, Hawthorne and Stanley make room for the intuition that probability claims can sometimes act as motivating reasons; in this case, RKP demands that we know the corresponding probabilities, which are further construed as epistemic, in an objective way. But we have no indications as to whether such objective measures can be defined and actually are known. The authors try to circumvent this problem by focusing on our knowledge of nonprobabilistic facts. However, just mentioning K is not enough—we still need explicit references to the appropriate measures, on pain of making a discourse about reasons unintelligible. So the problem remains. Moreover, it is not obvious that the linguistic evidence excludes a different, more subjective interpretation of ordinary probabilistic discourse, in which case resorting to knowledge/belief talk disregards crucial phenomenological aspects of the situation, as seen from the agent's point of view. In short, regardless of the merits of RKP for full beliefs, the attempt to squeeze all probability references (at the time of giving reasons) into RKP does not seem to be successful. The present reflections point to the fact that RKP cannot be the whole story on the link between knowledge and reasons for acting.

ELEONORA CRESTO

CONICET (Argentina)

MEMO TO AUTHORS

The Journal of Philosophy is published monthly, in issues that (ordinarily) contain only 56 pages of copy. This limits the length of papers normally published to no more than 7,500 words, or 18 *Journal* pages. A more precise way of determining length is by characters (including spaces): 2,850 characters per *Journal* page, or 51,300 characters per *Journal* article of 18 pages.

While the *Journal* does not publish manuscripts that mainly comment on articles published elsewhere, we do publish short responses to *Journal* articles. Submissions to this "Comments and Criticism" section should be no more than 5,000 words, or 12 *Journal* pages. *The Journal of Philosophy* does not publish papers that are primarily historical or expository, nor does it accept unsolicited book reviews.

We now require that all submissions be made by sending an email to submissions@journalofphilosophy.org with the manuscript file attached in PDF or Word format. Please include your name, contact information, and the title of your paper in your message. All submissions will be acknowledged by email, and all subsequent correspondence regarding the manuscript will be sent via email rather than post.

Further, manuscripts must be prepared for double-blind refereeing, with any and all revealing references to the author removed, including personal acknowledgments. Any submission that does not adequately conceal the identity of its author will not be read.

Please also observe the following guidelines when preparing submissions:

(1) All material, including quotations and notes, must be typed and double-spaced, with ample margins.
(2) No bibliographies will be published; bibliographical material should be put into notes.
(3) Notes should be complete, according to *Journal* style (see any issue). Journal citations should include: the first and last name of author(s), name of journal in full, volume number, issue number, month or season, year, the article's full range of pages, and, if applicable, the specified page or pages. Book references should include: the first and last name of author, publisher, U.S. city of publication, and date. Notes should be gathered at the end of the article.
(4) In quotations, spelling, capitalization, and formatting should adhere to the original.
(5) Special symbols should be clearly marked and entered in a font that complies with MS Word, such as Cambria Math.
(6) No underlining; please only italicize variables and the first appearance of technical terms.
(7) The use of contractions should be limited to questions.

The editors are willing to read and evaluate a manuscript that does not meet these seven guidelines, provided the author agrees to meet them before submitting a final manuscript. Please note that the editors may ask for revision of excessively long submissions and that the final version of accepted manuscripts must be formatted as a Word file.