



**UNIVERSITY OF
ILLINOIS PRESS**

The Possibility of Conceptual Clarity in Philosophy

Author(s): Michael A. Bishop

Source: *American Philosophical Quarterly*, Jul., 1992, Vol. 29, No. 3 (Jul., 1992), pp. 267-277

Published by: University of Illinois Press on behalf of the North American Philosophical Publications

Stable URL: <https://www.jstor.org/stable/20014420>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



and University of Illinois Press are collaborating with JSTOR to digitize, preserve and extend access to *American Philosophical Quarterly*

JSTOR

THE POSSIBILITY OF CONCEPTUAL CLARITY IN PHILOSOPHY

Michael A. Bishop

THE attempt to analyze and clarify concepts is a trademark of Western philosophy. And this is how it should be. Given the relatively non-empirical nature of the philosophical endeavor, philosophers must be concerned with the state of their primary instruments—language and the concepts expressed by language. The aim of this paper is to make sense of a pattern of argumentation typically employed in the effort to clarify philosophically important concepts. The upshot will be that only by adopting a thoroughgoing *naturalism* will we have any real chance to achieve the goal of conceptual clarity in philosophy.

I. COUNTEREXAMPLE PHILOSOPHY

Counterexample philosophy is a distinctive pattern of argumentation philosophers since Plato have employed when attempting to hone their conceptual tools. It begins when someone proposes a classical account of a philosophically important concept. A classical account of a concept offers singly necessary and jointly sufficient conditions for the application of a term expressing that concept. Probably the best known of these is the traditional account of knowledge.

(K) *X* is knowledge iff *X* is a justified true belief

The list of philosophers who have advanced classical accounts of concepts would not only include many of the greatest figures in the history of philosophy, but also many highly regarded contemporary philosophers. Consider just a fraction of these: David Lewis on convention (1969) and on causation (1973), Alvin Goldman on knowledge (1967), H.P. Grice on meaning (1957), Carl Hempel on explanation (1948), Wesley Salmon on causation (1984), and Hilary Putnam on mental states (1967).

The next stage of counterexample philosophy involves philosophical sharpshooters taking aim at the classical account of the concept by proposing counterexamples. While the paradigm is probably Gettier's (1963) counterexample to (K), counterexamples have also been levelled against each of the classical accounts noted above: Tyler Burge against Lewis' account of convention (1975), Jaegwon Kim against Lewis' account of causation (1973), Hilary Putnam against Goldman's account of knowledge (1983), Paul Ziff against Grice's account of meaning (1967), Sylvain Bromberger against Hempel's account of explanation (1966), Philip Kitcher against Salmon's account of causation (1989), and John Searle against Putnam's account of mental states (1980).

The final stage of counterexample philosophy involves one of three potential responses to an apparently successful counterexample. One might amend the account so it no longer succumbs to the counterexample, at which point philosophical sharpshooters reset their sights on the revised account. Or one might admit defeat: no amount of tinkering will save the account. Finally, one might bite the bullet by holding fast to the classical account even though many think the counterexample is successful. J. J. C. Smart's discussion of alleged counterexamples to act utilitarianism is a classic case of bullet biting: "The fact that [utilitarianism] has consequences which conflict with some of our particular moral judgements need not be decisive against it.... I am not concerned with what our moral customs and institutions in fact are... I am concerned with a certain view about what they *ought* to be" (1973, p. 56).

Which of these responses is appropriate depends on the philosopher's aim in adducing the account. Counterexample philosophy has three potential aims.

(1) *Conceptual analysis*: A serious and venerable psychological hypothesis holds that many concepts are made up of other concepts. For example, the concept *bachelor* might be constructed out of the concepts *available unmarried man*. If we adopt this psychological hypothesis, one might propose that the success condition on a classical account is that the concept expressed by the term on one flank of the biconditional be constituted by the concepts expressed by the terms on the other flank of the biconditional. For instance, (K) is a successful instance of *conceptual analysis* just in case our concept *knowledge* is made up of the conjunction of our concepts *justified true belief*.

(2) *Conceptual explication*: In order for an instance of *conceptual explication* to succeed, the extensions of the target term and its classical account must be identical with respect to all clear instances and non-instances of the target term.¹ On this view, (K) is a successful explication just in case anything that is a clear case of knowledge is also a clear case of justified true belief, and anything that is a clear non-instance of knowledge is also a clear non-instance of justified true belief. How our concepts are psychologically represented is irrelevant to this endeavor.

Sections II and III of this paper will offer empirical reasons why conceptual analysis and explication cannot succeed. If this is right, how are we to understand the attempt by some very distinguished philosophers to give classical accounts of philosophically important terms? Is it merely a mistake of a bygone era? Perhaps. But before we dispose of such a significant piece of philosophical heritage, we should consider the third potential aim of counterexample philosophy.

(3) *Conceptual revision*: So far we have considered conservative aims of counterexample philosophy—they set out to preserve something important about the target term. Conceptual revision does not set the preservation of the target term's intension or extension as a condition for success. Finding a difference between the extensions of the re-

vision and the target term is, by itself, irrelevant to the acceptability of the revision. But what makes one revision more successful than another? If there are no rules about this, if willy-nilly conceptual revision is the aim, then it seems a worthless and absurd project. Assuming the conservative endeavors (analysis and explication) are unlikely to succeed, it is essential we respond to this challenge on pain of forfeiting counterexample philosophy as a tool for conceptual clarification. Section IV is a response to this challenge.

One might argue that these options do not exhaust the potential goals of counterexample philosophy. Philosophers might set themselves the goal of providing an account of metaphysical reality. So (K) succeeds just in case it tells us what knowledge *really is*. Will the proponent of metaphysical counterexample philosophy offer an account of how we can know when we have found metaphysical reality? If not, then counterexample philosophy is futile since we cannot know when it succeeds. But suppose he does offer such an account. If part of the test for determining whether we have found metaphysical reality involves checking whether the extension of the target term has been preserved, at least in the clear cases, then metaphysical counterexample philosophy will fall victim to the argument against conceptual explication in section III. If metaphysical counterexample philosophy does not involve preserving the extension of the target term, then it faces the challenge faced by conceptual revision discussed above and in section IV. So for the purposes of this paper, metaphysical counterexample philosophy can be seen as either futile, an instance of explication, or an instance of revision.

II. THE FUTILITY OF CONCEPTUAL ANALYSIS

Classical conceptual analysis aims to describe the structure of a philosophically important concept in terms of singly necessary and jointly sufficient conditions. This endeavor is doomed because most concepts are not structured classically. This is not news to philosophers. Since Wittgenstein (1953), philosophers have argued that there are no classical conditions for the application of our

terms. But *even if* we assume that philosophically important concepts are represented classically, it does not follow that counterexample philosophy as currently practiced is a reasonable method to use to discover the structure of our concepts. Questions about the actual structure of our concepts are *in principle* as empirical as questions about the actual structure of iron. No one would seriously propose studying the structure of iron by engaging in counterexample philosophy. But one might suppose that counterexample philosophy plays an important supplementary role in the empirical study of our concepts. (For a necessarily abbreviated description of such studies, see the Appendix.)

Philosophical reflections on the nature of mental representation have influenced and should continue to influence the empirical study of concepts. But what can classical counterexample philosophy contribute to the empirical endeavor that cannot be attained more efficiently by psychologists? It has nothing to contribute about what concept C's structure might be (everyone knows it *might* be classical) or about what empirical methods to employ in finding out about C's structure. And considering the data psychologists take to be relevant to determining C's structure—which features subjects take to be defining, categorization times, typicality judgements, etc.—counterexample philosophy offers no evidence that psychologists can't readily discover on their own.

There is one situation in which counterexample philosophy could prove useful. It could be used to discover C's defining features when subjects are incapable of listing them, which could well happen with philosophically important terms. Two problems arise with this suggestion. First, until subjects fail to list defining features for a concept, there is no reason for the psychologist to engage in counterexample philosophy. So philosophers attempting to supplement psychology are jumping the gun. Second, and more importantly, since there is no *a priori* reason to assume that C is represented classically, there is no reason to search for *classical* conditions for the application of C. Thus, even when subjects can't deliver C's defining features, classical counterexample

philosophy is not the method that ought to be employed in finding those features.

Nicholas Rescher (1973, 1985) has advanced another argument to the conclusion that *even if* concepts are classically structured, conceptual analysis as practiced cannot succeed. Philosophically pivotal concepts, like most of our concepts, are "attuned to our practical dealings in a complex world where some degree of oversimplification is always necessary in the interests of manageability" (1985, p. 56). As a result, the set of criteria associated with such concepts inevitably reflect contingent assumptions (they are "fact coordinated"). For example, the concept of personal identity "unites a plurality of factors, among which bodily continuity and sameness of personality are the outstanding members. These factors are held together in a harmonious symbiosis by *factual* considerations" (1985, p. 50). As long as those assumptions are stable, the concept can be a productive unit of thought. But if those assumptions are sabotaged, paradox inevitably ensues. And those assumptions are constantly sabotaged when philosophers attempt to apply their analyses to very different possible worlds. When philosophers draw modal conclusions from their analyses or propose counterexamples that hold only in very distant possible worlds, they are undermining the conditions of the concept's effectiveness. Even if a concept is structured classically, its domain of effectiveness is limited to normal cases, and so applying its criteria to certain types of cleverly constructed cases will result in the sorts of paradoxes and counterexamples that pervade the philosophical literature.

Counterexample philosophy in the service of conceptual analysis is a futile endeavor. If philosophically important concepts are represented classically, counterexample philosophy as actually practiced is an irrational means of achieving the goal of conceptual analysis. And if they are not represented classically, the endeavor cannot succeed.

III. THE POSSIBILITY OF CONCEPTUAL EXPLICATION

Let's assume that the psychological literature is correct and our judgements about

whether a philosophically important term applies in actual and counterfactual situations are fixed by non-classical descriptions.² It is unlikely that we can find a non-circular *classical* explication of our terms.³ To see why, consider the situation. We begin with a set of objects and a language that has a fixed number of legitimate predicates. The extension of a predicate is a set of objects. The target term, P, is defined by some *disjunction* of predicates, A1-An. The extension of P is the union of the extensions of A1-An. What is the likelihood that there is a *conjunction* of predicates, B1-Bn, that is coextensive with P? At a purely intuitive level, it seems unlikely that there will be some set of predicates with extensions whose *intersection* is coextensive with P (the *union* of A1-An).

To cash out the intuition, consider what factors would tend to affect the likelihood of successfully explicating the empirical predicate, P. Here's one: All things being equal, it is more likely that we can find a classical explication for P insofar as the extension of P is smaller. The argument for this principle rests on two premises.

(1) Other things being equal, a smaller number of objects will fall within the extension of more empirical predicates of the language than a larger number of objects.

For example, the set of crows falls within the extension of more English predicates than the set of birds. This premise applies to empirical terms, but not to mathematical terms—they can be given classical explications even though they have infinite extensions. This paper assumes that the terms philosophers analyze apply to states of affairs that do not have the well-ordered structure of mathematics. So insofar as (1) is concerned, the terms philosophers analyze are more like empirical terms than mathematical ones. And surely the burden of proof rests on anyone claiming the opposite. The second premise is as follows.

(2) Other things being equal, the more singly necessary conditions a set of objects share, the more likely it is that some set of those conditions will be jointly sufficient.⁴

Given these premises, the probability of successfully explicating P is inversely related to the number of objects in its extension. As

we have seen, philosophical explications are meant to range over all logically possible worlds. Therefore, the chances of conceptual explication succeeding appears very low. But this conclusion comes too fast. There are other factors that affect the probability of success in conceptual explication, and some of these might counteract the extension effect. We have indirect evidence suggesting there is no countervailing factor: Philosophers since Plato have failed to propose a classical explication of a philosophically important term that is today considered successful. So we may legitimately conclude that all of the evidence available to us about the situation facing those engaged in classical explication indicates that it is very likely to fail.

IV. CONCEPTUAL REVISION

If the attempt to offer accounts of philosophically important concepts that preserve the original term's extension is futile, then what becomes of philosophy? Should we bequeath the attempt to analyze and understand our concepts to empirical psychology? Perhaps. But before doing this, let's consider the option of conceptual revision. The revisionist does not set for herself the goal of preserving the extension of a term as it is used by a group of speakers. But then, what makes one potential revision preferable to another? One plausible variety of conceptual revision involves *the principle of theoretical excellence*: Adopt the successor concept that is part of the best available empirical theory. The 'best available empirical theory' is the one that possesses the greatest combination of theoretical virtues (such as explanatory and predictive power, accuracy, simplicity, scope).

The motives behind the adoption of this principle are the same ones we have for accepting the best available scientific theories. Although there is controversy on this point, there are two reasons to adopt the aim of conceptual revision in the service of theoretical excellence. First, the best available theories usually provide pragmatic benefits over less adequate theories—individuals and societies tend to be more efficient in achieving their goals and meeting their needs as their

theories become more pragmatically powerful. Second, one is more likely to have true beliefs if one adopts the best available theory.⁵ So if the extension of a successor concept is reasonably similar to that of the target concept, and the successor plays a role in a theory that provides more pragmatic and epistemic benefits than any alternative, then it is reasonable to adopt it as a revision of the target concept. Is conceptual revision in the service of theoretical excellence an achievable goal? At least in some cases, there is reason to think it is.

IV.i. *Science: The Future of Philosophy?*

Bertrand Russell believed that many philosophical issues spun off from philosophy as soon as they became well understood. This characteristic centrifugal force accounts in part for the uncertainty of philosophy.

[A]s soon as definite knowledge concerning any subject becomes possible, this subject ceases to be called philosophy, and becomes a separate science. The whole study of the heavens, which now belongs to astronomy, was once included in philosophy; Newton's great work was called "the mathematical principles of natural philosophy"... Thus, to a great extent, the uncertainty of philosophy is more apparent than real: those questions which are already capable of definite answers are placed in the sciences, while those only to which, at present, no definite answer can be given, remain to form the residue which is called philosophy (1959, p. 155).

Science contains numerous well-defined concepts and even some classical ones. Of course, many scientific concepts have no strictly defined conditions of application and yet are used to good effect. This is especially so in the "higher-level" sciences like biology, psychology and the social sciences. Whether this conceptual fuzziness is a permanent or essential feature of these sciences a very hard and tricky issue. Nevertheless, the history of science gives us some reason to believe that as a science matures, the concepts it employs become clearer.

Chemistry is an example of a science that contains many classical and well-defined concepts. Some of these (e.g., atom and water) have been the object of considerable

philosophical contemplation. If we look closely at how such terms are treated by science, however, we quickly realize that the concepts and their extensions have been radically revised. Consider the term "gold." Chemical theory identifies gold as the stuff with a particular chemical structure, while most people are disposed to apply "gold" to things that satisfy descriptions like " x is yellowish, x is valuable, x is a metal..." If this is right, it is easy to develop counterexamples to the chemical account of the term "gold"—counterexamples that show that the extension of the chemical concept differs from the extension of our commonsense concept. For example, suppose that in some possible world the element with atomic number 79 were a purple, viscous liquid. Many would not be disposed to apply the term "gold" to this purple goo. Suppose also that in this possible world, there is a metal that has all the aesthetic and economic properties gold has in the actual world, but its atomic number isn't 79. Many people would be disposed to apply the term "gold" to this. The accounts of concepts given by scientific theories are often *revisions*, sometimes radical revisions, of the concepts we actually employ and their extensions across possible worlds.⁶

One might argue that the chemical account is not really a revision since the extension of "gold" has always been the stuff with atomic number 79, regardless of how people are disposed to apply the term in counterfactual situations. Whether or not this account of the extension of natural kind terms is correct, the chemical account is revisionary in the following sense: It is not the account that would be acceptable to those engaged in counterexample philosophy before they had adopted chemical theory. For them, the sole test for the acceptability of a proposed account of the term "gold" is how they are disposed to apply it in counterfactual situations. Insofar as we are interested in the product of counterexample philosophy, it is legitimate to say that the chemical account of "gold" is revisionary.

Suppose a concept which is deemed to be philosophically important today, will be appropriated by scientists when that concept is fit for empirical study. One way philosophers can succeed in classically revising their con-

cepts is to ready them for appropriation. Scientists will find a concept useful when it fits into a promising empirical theory. Therefore, one way philosophers can succeed in classically revising, or at least clarifying, their concepts is by attempting to fit them into empirical theories. It is unquestionably perverse to suggest that philosophers can succeed in classically revising their concepts by handing them over to scientists. But the world is sometimes perverse. Further, if Russell is correct, philosophers have no choice in the matter: part of their *raison d'être* is to bequeath useful concepts to science.

This sort of conceptual transfer is a real possibility in the philosophy of science. Concepts that have been of particular concern in this arena, such as scientific theory, explanation, confirmation, pseudo-science, etc., may (or may not) eventually play a role in empirical attempts to understand the process of science. From what we've seen, it is unlikely that we will ever find classical accounts of these terms that will also preserve their current extensions. But if similar successor concepts play a role in the history, sociology, or psychology of science, there is at least some hope that some day they will be defined clearly or perhaps even classically.

IV.ii. *When Science is not the Goal*

In an earlier passage, Russell claimed that when an area of philosophy becomes understood, it spins off to form a science. He has more to say on this subject.

This is, however, only part of the truth concerning the uncertainty of philosophy. There are many questions—and among them those that are of the profoundest interest to our spiritual life—which, so far as we can see, must remain insoluble to the human intellect unless its powers become of quite a different order from what they are now... Yet, however slight may be the hope of discovering an answer, it is part of the business of philosophy to continue the consideration of such questions, to make us aware of their importance, to examine all the approaches to them, and to keep alive that speculative interest in the universe which is apt to be killed by confining ourselves to definitely ascertainable knowledge (1959, p. 155-6).

Philosophers have long been concerned with critical issues that appear to have no place in empirical or explanatory theories. These concepts are not amenable to conceptual revision in the service of theoretical excellence. It makes no sense to adopt that revision which plays a role in the best available empirical theory if the original has no role whatsoever to play in such a theory.

What should we conclude about the possibility of clarifying philosophically important concepts that play no role in empirical theories? Given the best psychological data available, we probably cannot find a classical account of such concepts that preserve their intensions or extensions. So we have a choice. We can leave to empirical psychology the attempt to plumb the structure of these concepts or we can embrace conceptual revision. But again we face the problem of what makes one revision better than another. The proper solution is essentially the same as before: Adopt that revision that plays a role in the best available *non-empirical* theory on the subject. Given that philosophers are often not concerned with definitely ascertainable knowledge, we need not insist that the best theories available about some subject must be *scientific* theories.

Does a revisionary counterexample philosophy that seeks non-empirical theoretical excellence have some possibility of success? In science, the attempt to construct integrated, coherent theories over time tends to produce well-defined, and sometimes classical, concepts. Perhaps the same results can be achieved in philosophy if the same method—theory construction—is employed. This argument from analogy is by no means apodictic. Nor should we expect it to be since philosophy seldom offers guarantees of success. But at least the empirical evidence does not doom this revisionary endeavor (unlike its more conservative counterparts).

How are philosophers supposed to proceed in constructing non-empirical philosophical theories? John Rawls offers some clear advice concerning theories of justice.

I wish to stress that a theory of justice is precisely that, namely, a theory. It is a theory of the moral sentiments (to recall an eighteenth cen-

tury title) setting out the principles governing our moral powers, or, more specifically, our sense of justice. There is a definite if limited class of facts against which conjectured principles can be checked, namely, our considered judgments in reflective equilibrium. A theory of justice is subject to the same rules of method as other theories. Definitions and analyses of meaning do not have a special place: definition is but one device used in setting up the general structure of theory. Once the whole framework is worked out, definitions have no distinct status and stand or fall with the theory itself... This is the conception of the subject adopted by most classical British writers through Sidgwick. I see no reason to depart from it. [In a footnote] I believe this view goes back in its essentials to Aristotle's procedure in the *Nicomachean Ethics* (1971, p. 50-1).

We can identify three conditions that should be met by anyone proposing to clarify philosophically important concepts that appear to have no role to play in empirical theories. First, the philosopher needs to identify the phenomena to be accounted for by the theory. In the course of theory construction, his conception of what the theory should account for might change. But at any time, he should be clear about his object of study. Second, the philosopher should attempt to provide some sort of explanatory or theoretical structure to account for the relevant phenomena. Definitions are helpful only insofar as they aid in the production or in the understanding of a theory, but they are not the primary aim of the enterprise. And third, theory-choice should be governed by the same sorts of normative principles as are empirical theories. Much more needs to be said about the nature of non-empirical theory construction in philosophy. But these three conditions are important because they can be (and have been) so readily overlooked in the philosophical attempt to clarify our concepts.

IV.iii. *Counterexample Philosophy and Conceptual Revision*

Counterexample philosophy is particularly efficient at exposing the ways in which the extensions of two concepts differ. However, if we want to produce a revision of an old concept, what role, if any, does counter-

example philosophy have to play? It retains an important role. Engaging in conceptual revision does not guarantee freedom from counterexamples. In order for a counterexample to succeed against a revision, however, it can't merely show that the new concept has a different extension from the original. It must show that the revised concept fails to perform its appointed role within the theory in which it is embedded. The trick to engaging in revisionary counterexample philosophy is to begin by understanding the relevant theory and the role the target concept is supposed to play in that theory. In this way, we learn to properly employ that concept. Once we have grasped the new concept, it is possible to proceed exactly as if we were engaged in conceptual explication. That is, we check the extension of the proposed account against the extension of the new concept as we understand it. The only difference between conceptual explication and conceptual revision is that the latter doesn't explicate a concept already in use. It explicates a concept that is part of a developing theory. (In at least some cases, it is likely that the theory itself will provide the classical conditions for the application of the term. For example, chemical theory provides the resources for a classical account of the term "water.")

V. FINAL WORDS

This paper has been concerned with the philosophical attempt to analyze and clarify philosophically important concepts. The general moral to draw is that this philosophical endeavor, and probably lots of others, are hostage to empirical facts. Without the assumption that concepts are represented classically, classical accounts that aim to preserve the intension or extension of a term don't make a whole lot of sense. A second conclusion to draw is that if we want counterexample philosophy to succeed, we need to alter its character by making it more naturalistic. This involves viewing it as either part of an attempt to construct an empirical theory or as part of an attempt to construct a non-empirical theory that is subject to some of the same demands as empirical theories. In particular, it should have a reasonably well-

defined range of data it is supposed to account for, it should provide some sort of explanatory structure for that data, and it should be subject to the same sorts of normative principles as are empirical theories.

Making counterexample philosophy more naturalistic both increases the probability of finding a classical account of a concept and carries with it the benefits associated with adopting a successful theory.⁷

Iowa State University

Received October 28, 1991

APPENDIX: THE PSYCHOLOGICAL DATA

Over the past twenty years or so, psychologists have adduced considerable evidence for thinking that concepts are not structured in terms of singly necessary and jointly sufficient conditions. Psychologists studying the structure of our concepts hope to understand the cognitive structure and processing of natural language. They construct cognitive models of concepts that are faithful to data about the ways in which we employ our concepts, primarily in categorization tasks. Most of these models assume that a concept is instantiated as an interpretable syntactic string (or a set of such strings). It is natural and conventional to call these interpreted syntactic strings *descriptions*. The classical view of concepts (which is reminiscent of the classical description theory of terms) says that the descriptions that constitute concept *C* provide singly necessary and jointly sufficient conditions for the application of *C*. So in order for an object *O* to be an instance of *C*, *O* must satisfy all the descriptions that constitute *C*.

In order to make behavioral predictions about how concepts are actually employed, we need more than just an account of their structure. We need a *model* of our concepts that describes the cognitive mechanisms subserving particular tasks, such as categorization. Consider the following (simple and incomplete) classical model of concepts.

- (I) Concepts consist of sets of descriptions that are singly necessary and jointly sufficient for the application of the concept.
- (II) If C_2 is a superordinate of C_1 , then the descriptions that constitute C_2 also constitute C_1 .
- (III) Categorization works as follows: The comparator takes each description that constitutes C_2 (the alleged superordinate, say, bird) and searches C_1 (the alleged subordinate, say, robin) for a match. If all of C_2 's descriptions are matched, then C_1 is categorized as a C_2 . If, however, any of C_2 's descriptions do not match one of C_1 's descriptions, then C_1 is not a C_2 .
- (IV) The time it takes to complete the matching process is an increasing function of the number of description-description matches that are attempted.

(I) is the heart of the classical model. To give it up is to give up the assumption that concepts are structured classically. (II)—(IV) describe the mechanisms subserving categorization. They explain how a subject decides that a robin is a bird, or that a grasshopper is not a fish, or that the *belief that the man who will get the job has ten coins in his pocket* is (or is not) knowledge.

The problem with the classical view of concepts is that no motivated classical model has been constructed that can account for *typicality effects* while competing models account for them quite well (for a survey, see Smith and Medin, 1981). Consider the first piece of evidence marshalled against the classical view.

- (1) *Non-necessary features listed as defining*: When subjects are asked to give descriptions of the defining features of a category, they often list non-necessary features (Hampton, 1979). For example, when given the category *vehicle*, most subjects offered the non-necessary feature "has wheels" as a defining feature (counterexamples: boats, sleds).

The fact that subjects list non-necessary features as defining is not by itself telling evidence against the classical view. It is possible that subjects report features typically associated with vehicles that are not constitutive of their concept *vehicle*.

- (2) *Determinant of Typicality*: C_1 tends to be ranked as more typical of C_2 to the extent C_1 shares defining features with C_2 . For example, *robin* will be ranked as a more typical *bird* than (say) *chicken* because it shares more defining features with *bird* than does *chicken* (including the particularly important feature, “flies”) (Rosch and Mervis 1975; Hampton 1979).

Again, this need not shake the faith of the defender of the classical model. The model is silent on how typicality judgements are made. (Ideally, of course, a classical model will eventually be able to explain them.)

- (3) *Typicality Effect*: If C_1 is more typical than C_1' of C_2 , then C_1 will be categorized *faster* than C_1' as an instance of C_2 . So *robin* will be categorized faster than *chicken* as an instance of *bird* (Rosch, 1973; Hampton, 1979).

Given the classical model, why do subjects categorize *robin* faster than *chicken* as an instance of *bird*? Presumably something about the concept *chicken* makes it harder to categorize than *robin*. Perhaps the descriptions that constitute the concept *chicken* are harder to retrieve. Or perhaps it consists of more descriptions than does the concept *robin*. After all, the classical model implies that when a subject is asked whether a C_1 is a C_2 , assuming the answer is “Yes,” categorization time will be an increasing function of the number of description-description matchings that must be carried out. Given the data adduced in 1–3, the classical model is not yet in danger.

- (4) *Typicality Effects—A Function of the Superordinate*: Whether C_1 is categorized faster than C_1' as an instance of C_2 depends on C_2 . For example, while *robin* is categorized faster than *chicken* as an instance of *bird*, *chicken* is categorized faster than *robin* as an instance of *animal* (Smith and Medin, p. 51-3).

Now the classical model is in trouble. Categorization times are a function of the superordinate concept involved, not the subordinate concepts. Assuming that there is something about the concept *chicken* that makes it harder to categorize than *robin* flies in the face of the fact that *chicken* is categorized faster than *robin* as an instance of *animal*. How then to account for the typicality effects on the classical view of concepts?

The simple classical model also has problems with nested concepts. A trio of concepts are nested when C_2 is a superordinate of C_1 , and C_3 is a superordinate of C_2 . For example, the trio *robin*-*bird*-*animal* are nested. The classical model predicts that C_1 will be categorized faster as a C_3 than as a C_2 , and this is false for most cases (Smith & Medin, 1981).

Of course, it is possible to tinker with the simple model so it does not succumb to these problems. For instance, by giving up (II), we might be able to save the classical assumption and avoid the problems that arise with the typicality effects. However, giving up (II) makes it very difficult to explain how categorization might work. How can C_1 be categorized as an instance of C_2 if C_1 does not consist of all the descriptions that are necessary and sufficient for the application of C_2 ? No classical models yet proposed seem capable of accounting for the above categorization phenomena.

The failure of classical models is compounded by the fact that alternative models do a much better job handling the available evidence. Consider a probabilistic model of concepts (which is reminiscent of the cluster theory of names [Wittgenstein, 1953; Searle, 1958]): Concepts consist of a set of descriptions that need not be individually necessary for the application of the concept. Each description is given a weight, which is a measure of the subjective probability that instances of the concept will possess the property described. Categorization is (again) a matching process, but when a match occurs, the matched description's weight is added to a running sum. C is categorized as an instance of C_1 just in case C possesses a critical sum of the weighted features. Behavioral results that proved problematic for the classical model are easily explained on the probabilistic model. For example, the typicality effect is the result of the critical sum of weighted descriptions being arrived at quickly. Many psychologists believe that the probabilistic model is inadequate. For this reason, new models of concepts (particularly exemplar or prototype models) have appeared (Smith and Medin, chapter 7).

BIBLIOGRAPHY

- Bromberger, Sylvain. (1966), "Why-Questions," in Robert G. Colodny, ed., *Mind and Cosmos*. Pittsburgh: University of Pittsburgh Press.
- Burge, Tyler. (1975), "On Knowledge and Convention," in *The Philosophical Review*, vol. 84, pp. 249-55.
- Carnap, Rudolf. (1956), *Meaning and Necessity*. University of Chicago Press.
- Gettier, E.L. (1963), "Is Justified True Belief Knowledge?," in *Analysis*, vol. 23, pp. 121-23.
- Goldman, Alvin. (1967), "A Causal Theory of Knowing," in *Journal of Philosophy*, vol. 64, pp. 353-73.
- Grice, H.P. (1957), "Meaning," in *The Philosophical Review*, vol. 66, pp. 377-88.
- Hampton, J. (1979), "Polymorphous Concepts in Semantic Memory," in *Journal of Verbal Learning and Verbal Behavior*, vol. 18, pp. 441-61.
- Hempel, Carl. (1948), "Studies in the Logic of Explanation," in *Philosophy of Science*, vol. 15, pp. 135-75.
- Kim, Jaegwon. (1973), "Causes and Counterfactuals," in *The Journal of Philosophy*, vol. 70, pp. 570-72.
- Kitcher, Philip. (1989), "Explanatory Unification and the Causal Structure of the World," in Kitcher and Salmon, eds., *Minnesota Studies in the Philosophy of Science*, vol. 13, pp. 410-505.
- Lewis, David. (1969), *Convention: A Philosophical Study*. Cambridge: Harvard University Press.
- _____. (1973), "Causation," in *Journal of Philosophy*, vol. 70, pp. 556-67.
- Putnam, Hilary. (1967), "The Nature of Mental States," in W.H. Capitan and D.D. Merrill, eds., *Art, Mind, and Religion*. Pittsburgh: University of Pittsburgh Press.
- _____. (1975), "The Meaning of 'Meaning'," in *Language, Mind, and Knowledge*, ed. Keith Gunderson. Minneapolis: University of Minnesota Press.
- _____. (1983), "Why Reason Can't Be Naturalized," in Putnam, *Realism and Reason*. Cambridge: Cambridge University Press.
- Rawls, John. (1971), *A Theory of Justice*. Cambridge: Harvard University Press.
- Rescher, Nicholas. (1973), *The Primacy of Practice*. Oxford: Basil Blackwell.
- _____. (1985), *The Strife of Systems*. Pittsburgh: University of Pittsburgh Press.
- Rosch, E. (1973), "On the Internal Structure of Perceptual and Semantic Categories," in T.E. Moore, ed., *Cognitive Development and the Acquisition of Language*. New York: Academic Press.
- Rosch, E., and Mervis, C.B. (1975), "Family Resemblance Studies in the Internal Structure of Categories," in *Cognitive Psychology*, vol. 7, pp. 573-605.
- Russell, Bertrand. (1959), *Problems of Philosophy*. Oxford University Press.
- Salmon, Wesley. (1984), *Scientific Explanation and the Causal Structure of the World*. Princeton: Princeton University Press.
- Searle, John. (1958), "Proper Names," *Mind*, vol. 67, pp. 166-73.
- _____. (1980), "Minds, Brains, and Programs," in *Behavioral and Brain Sciences*, vol. 3, pp. 417-58.
- Smart, J.J.C. (1973), *Utilitarianism: For and Against*, with Bernard Williams. Cambridge: Cambridge University Press.
- Smith, Edward, and Medin, Douglas. (1981), *Categories and Concepts*. Cambridge: Harvard University Press.
- Wittgenstein, Ludwig. (1953), *Philosophical Investigations*, trans. G.E.M. Anscombe. Oxford: Basil Blackwell.
- Ziff, Paul. (1967), "On H.P. Grice's Account of Meaning," in *Analysis*, vol. 28, pp. 1-8.

NOTES

1. Carnap introduced the notion of *conceptual explication* as follows: One is engaged in conceptual explication if one is "making more exact a vague or not quite exact concept used in everyday life or in an earlier stage of scientific or logical development, or... replacing it by a newly constructed, more exact concept..." (1956, p. 7-8). The account of explication given here assumes that part of what is involved in providing a "more exact" concept is providing a classical concept.

2. Although it is plausible to assume that intension alone does not always fix extension, the assumption that it does is justified in this context because of the nature of counterexample philosophy. When a counterexample is given against an account of C, the success of the counterexample depends on whether subjects judge it to be an instance or non-instance of C. And these judgments are fixed solely by the terms' intensions. So in attempting to understand the practice of counterexample philosophy, it is legitimate to assume that intension fixes extension.
3. Some might argue that the intensions of some terms embody their "real essences." Thus, the intension of "water" is given by the description, "the stuff having the same chemical nature as *this*" (Putnam, 1975). Therefore, a successful explication will specify that chemical nature. The problem with this suggestion is that this cannot be an analysis of how our concepts are actually represented. Since individuals confidently apply the term "water" to a liquid without knowing its chemical composition, other descriptions (including perhaps the description involving the chemical properties of water) must be involved.
4. One might argue that this factor cuts both ways. Although it may be easier to find *necessary* conditions for a smaller number of objects, it is harder to find *sufficient* conditions. This objection is irrelevant. Given that conceptual explications range over a huge (or an infinite) number of objects, it is easy to find sufficient conditions for the application of a concept. However, a successful *classical* explication will not simply provide a disjunctive list of sufficient conditions. It will provide list of necessary conditions that are jointly sufficient. And a larger set of objects to account for makes only the former task easier.
5. Certain antirealists will not be moved by this second motive. For them, the pragmatic motive is sufficient to adopt the aim of conceptual revision in the service of theoretical excellence.
6. One might argue that everyone is disposed to apply "gold" in the counterfactual examples in accordance with the chemical account. But this is doubtful. No one would insist upon a chemical account of terms like "mud" or "tabasco sauce." Why would we expect the chemically uninitiated to treat the mass noun "mud" differently from the mass noun "water"? Ultimately, of course, this is an empirical issue and requires empirical investigation.
7. I would like to thank Patricia Kitcher, Philip Kitcher, David Magnus, the members of the philosophy department at Iowa State University, the editor and an anonymous referee for very helpful comments on earlier drafts of this paper. I owe special thanks to Stephen Stich for many discussions and ideas.