



Concepts as Tools in the Experimental Generation of Knowledge in Cognitive Neuropsychology

Author(s): Uljana Feest

Source: *Spontaneous Generations: A Journal for the History and Philosophy of Science*, Vol. 4, No. 1 (2010) 173-190.

Published by: The University of Toronto

DOI: [10.4245/sponge.v4i1.11938](https://doi.org/10.4245/sponge.v4i1.11938)

EDITORIAL OFFICES

Institute for the History and Philosophy of Science and Technology
Room 316 Victoria College, 91 Charles Street West
Toronto, Ontario, Canada M5S 1K7
hpsat.society@utoronto.ca

Published online at jps.library.utoronto.ca/index.php/SpontaneousGenerations
ISSN 1913 0465

Founded in 2006, *Spontaneous Generations* is an online academic journal published by graduate students at the Institute for the History and Philosophy of Science and Technology, University of Toronto. There is no subscription or membership fee. *Spontaneous Generations* provides immediate open access to its content on the principle that making research freely available to the public supports a greater global exchange of knowledge.

Concepts as Tools in the Experimental Generation of Knowledge in Cognitive Neuropsychology*

Uljana Feest[†]

This paper asks (a) how new scientific objects of research are conceptualized at a point in time when little is known about them, and (b) how those conceptualizations, in turn, figure in the process of investigating the phenomena in question. Contrasting my approach with existing notions of concepts and situating it in relation to existing discussions about the epistemology of experimentation, I propose to think of concepts as research tools. I elaborate on the conception of a tool that informs my account. Narrowing my focus to phenomena in cognitive neuropsychology, I then illustrate my thesis with the example of the concept of implicit memory. This account is based on an original reconstruction of the nature and function of operationism in psychology.

I. INTRODUCTION

Let me start out by describing what I take to be the *prima facie* puzzle when talking about concepts in the context of scientific research: In order to investigate a given phenomenon, one has to be able to empirically individuate instances of it. In order to be able to do so, one has to possess some concept of the phenomenon. The possession of a concept is generally taken to imply knowledge about the class of phenomena that it applies to. But how do we make sense of concept use in cases where scientists investigate phenomena or objects they don't know much about, or perhaps aren't even sure really exist? Clearly they must use concepts when conducting their research. The question that interests me here is what concepts have to be like, such that they can play a role in the empirical generation of knowledge.

* Received 26 February 2010. Revised paper accepted 28 June 2010.

[†] Uljana Feest is an assistant professor at the Institut für Philosophie, Literatur-, Wissenschafts- und Technikgeschichte at the Technische Universität Berlin. Her research interests include the epistemology of experimentation and the philosophy and history of psychology and the human sciences.

In this paper I will suggest that we think of concepts as *tools* in the investigative process. They are tools for knowledge generation, and like other tools they can be adapted or discarded in the process. More specifically, I propose an account of *operational definitions* that explicates how concepts can fulfill this task, and I present an account of a recent episode in the history of cognitive psychology to demonstrate my thesis. My account of concepts as research tools differs from existing accounts of concepts by virtue of the fact that I ask a different question than they typically ask. In the philosophical literature, the question about the relationship between concepts and what we know about their referents is usually asked by either asking about the *semantics* of the words that correspond to the concepts, or by asking about the *epistemology* of finding out that an object really has the properties ascribed to it by a given concept. In contrast, I will argue that if we are interested in the role of concepts in knowledge acquisition, it is instructive to ask (1) not what a given concept means, but rather what scientists *take it to mean*, and how the scientists' taking-the-concept-to-mean-something can play a productive role in empirical research, and (2) not whether an already fully formed concept is really a good instrument for individuating an object, but how the concept is developed in the process of using it as an instrument.

Before I proceed to present my account, let me emphasize that I explicate my approach in relation to a couple of well known existing accounts of concepts mainly for expository purposes. It is *not* the point of this article to develop a general theory of concepts. To phrase this differently, it is not my aim to make a contribution to questions about meaning and reference that have been debated in the philosophy of language for a long time. Rather, I wish to draw attention to an entirely different set of questions—questions about experimental knowledge generation—which, I claim, can be elucidated by thinking about the role of concepts in experimental practice in the way proposed here.

II. THE CAUSAL THEORY OF REFERENCE AND THE INVESTIGATIVE PROCESS

In the literature on concepts, scholars sometimes distinguish between the “classical” theory of concepts and “non-classical” theories of concepts (e.g., Laurence and Margolis 1999). According to the classical theory, concepts have a definitional structure, which provides necessary and sufficient conditions of application for the words that correspond to the concepts. This definitional structure, furthermore, is frequently equated with what Frege termed the *sense* (as opposed to the *reference*) of the concept. Possession of a concept enables the subject to identify objects that are in the extension of the concept. The best-known instances of the

classical theory of concepts can be found in forms of empiricism, which are commonly held to assume that concepts can be defined in terms of observational features.

There are two groups of standard reasons for why the classical view failed, namely (1) counterexamples, which show that almost any definition in terms of necessary and sufficient conditions excludes cases we intuitively take to be in the extension of the concept, and vice versa, and (2) psychological evidence to the effect that concept formation/application does not work the way suggested by the classical theory (e.g., Rosch 1978). This latter type of work has in turn given rise to work that models conceptual change in science on cognitive theories of concepts (e.g., Andersen et al. 2006). Interesting as this latter work is, I shall not be paying attention to it here, as my interest does not lie with the cognitive mechanisms of developing and applying concepts, but with the *methodological role* of assuming that certain concepts apply in the context of scientific research. By “methodological role” I mean the role of enabling subjects to conduct research about purported objects, that is, objects whose existence and characteristics are still in question. To understand the desiderata of such an analysis it will be helpful to review some responses to the classical theory of concepts and its successors.

There are two well-known theoretical responses to the problem with the classical theory of concepts just outlined. Roughly, we may identify them as the description theory of meaning and the causal theory of reference, which are associated with the names of Willard Van Orman Quine and Hilary Putnam respectively. Each of these theories of meaning, I would like to suggest, explicates some ideas we have about scientific concepts, namely (in the case of the description theory) that they are revisable and (in the case of the causal theory) that they have stable referents, which do not change simply because our concepts of them change. However, both theories of meaning fall short of providing an analysis of how concepts can figure in the generation of knowledge. This need not necessarily be a shortcoming, since these theories do not aim at providing an account of knowledge generation. Given my own interest in this question, it is helpful to review in which ways they fail to answer it. I will start with the causal theory.

The best-known formulation of the causal theory was provided by Putnam in his 1975 “The Meaning of ‘Meaning.’” This paper takes as a point of departure the fact that scientists sometimes fundamentally change their most basic concept of a given object, and yet continue to talk as if they were still talking about the same object. A well-known instance of this is provided by the fact that non-Euclidean geometry radically altered the concept of physical space, yet scientists continued to assume that

the term “physical space” referred to the same phenomenon before and after this change. Putnam’s causal theory of reference is an attempt to justify this practice by saying that a word’s meaning gets fixed when it is first introduced into the language (based on the object that was present at the time). This ensures the idea of a stable referent of a word in the absence of a well thought-out concept of what it refers to. However, as has variously been pointed out, one problem with this account is that it is unable to predict what happens when people start using the term in a way that is either broader or narrower than before. The causal theory has to argue in both cases that people are either misapplying the concept, or that the concept had the broader/narrower extension all along. An example provided by Gary Ebbs (2003) makes this point: In the seventeenth century, people used the term “gold” to refer to both what we today refer to as gold and as platinum, based on observational similarities between the two substances. Once it was determined that they have different atomic numbers, the extension of the concept was narrowed down. As Ebbs argues, it is equally conceivable that people might simply have distinguished between two kinds of gold. The point is that there are no a priori philosophical arguments available to predict which way a particular linguistic community will develop their concepts: “what we *actually* say when we find ourselves in previously unimagined situations almost always trumps our earlier speculations about what we *would* say if we were to find ourselves in that situation” (Ebbs 2003, 252). He suggests that our practical judgment of sameness-of-denotation across time is much more robust than any of the strategies that people have appealed to in order to justify it: “we trust these judgments unless we have some special reason to doubt or revise them” (Ebbs 2003, 252).¹

I would like to adopt this proposal here. It makes room for the notion that scientists view their phenomena of research as relatively stable features of the world, in the sense of Bogen and Woodward (1988), but also that they can at times revise a concept to the point of giving up the very idea that the phenomenon even exists. In addition, I suggest that the *judgment of sameness of denotation* is a prerequisite for research on a given purported object. It is in this practical reliance on scientific judgments about the denotations of concepts that provides a first, minimal, explication of my contention that concepts can be *tools* in the generation of knowledge. But of course, so far nothing has been said about the question of how such tools are used. In order to get a clearer sense of that, we need to say more

¹ Ebbs’s criticism is directed both at older formulations of the causal theory of meaning and at more recent refined versions, such as the one presented in two-dimensional semantics (e.g., Chalmers 1996; Jackson 1998).

about empirical criteria of application for the concept. After all, the notion that scientists work with the assumption of a stable object of research, does not, in and of itself, give us any indication as to how they go about recognizing instances of it, let alone how they proceed in their attempts to find out more about it. In other words, we need to say more about the descriptive features specified by a concept.

III. OPERATIONAL DEFINITIONS AND THE EPISTEMOLOGY OF DISCOVERY

Since my focus here is on experimental investigations, I shall assume that the relevant descriptive features are a result of, and play a role in, experimental interventions. This focus on experimental interventions comes with a proposal to think of the descriptive features of a concept not in terms of whether they can adequately represent the object under investigation, but *how they enable experimental interventions in the process of investigating the purported or ill-understood object*. The basic idea here is that concepts figure as tools for the investigation of such objects. As such they can contribute to experimental knowledge generation, but they can also be refined and discarded in the process.

Operational Definitions and the Logic of Disposition Terms

I am proposing here a notion of *operational definition* as an analysis of how concepts fulfill this task. My understanding of operational definitions differs from the usual philosophical usage of the expression. There, it is often taken (and therefore rejected) as a particularly clear example of the classical theory of concepts. One well-known point of reference here is Percy Bridgman's remark that a "concept is synonymous with the corresponding set of operations" (Bridgman 1927, 5). Now, it is an open question whether Bridgman really intended his thesis to make a contribution to a theory of meaning in the sense of verificationist semantics. Rather, he wanted to caution scientists not to assume uncritically that two experiments really operationalize the same concepts, that is, pick out the same phenomenon.² However, even if we read Bridgman's thesis as one about semantics, it would have to be read as a thesis about the semantics of a specific class of terms, namely *disposition terms* ("if I were to perform operation x then observable event y would occur"). Such terms—as Carnap recognized by 1936—do not easily lend themselves to definitions by means of the classical theory of concepts. This prompted Carnap to propose an analysis of disposition terms, which

² See Chang (2009) for a discussion of Bridgman's operationalism.

did not provide necessary conditions of application. Instead he merely suggested to specify the meaning of a disposition term in terms of one specific condition of application (leaving open the possibility that it might also apply in others).

$Q_1 \rightarrow (Q_3 \text{ iff } Q_2)$ (Carnap 1936, 443)
(e.g., "If an object is put into water, then it is water-soluble iff it dissolves.")

My aim here is not to defend Carnap's semantic analysis of disposition terms, but rather to ask what experimental possibilities are opened up for scientists if they conceptualize the purported referent of a given word in a way that is analogous to this semantic analysis. Consider the following hypothetical research scenario: we want to investigate the molecular structure of water-soluble substances. In order to do so we need to identify the members of our test class empirically. We can do so by specifying an experimental set-up in which a sample of the substance is put in water to see if it dissolves. Such specifications may well involve instructions about the types of materials to use, as well as about the temperature of both the room and the water. Notice that while some such specification is a vital prerequisite for doing research on soluble substances, it is not infallible or unrevisable. For example, it may later turn out that under the specified conditions some water-soluble substances do not dissolve, or that some *appear* to, but do not in fact, dissolve.

Operational Definitions as Providing Instructions for Experimental Interventions

The basic point here is that we cannot even begin to study the purported object of research (in this case, the molecular structure of water-solubility) unless we work with a preliminary understanding of how to empirically individuate the objects that possess it. Operational definitions function as tools to this end by providing paradigmatic conditions of application for the concepts in question. These are cast in terms of a description of a typical experimental set-up thought to produce data that are indicative of the phenomenon picked out by the concept. I base my analysis of the role of operational definitions on several historical case studies where I have traced the origins of talk about operationism in psychology to psychophysics and behaviorism of the 1920s and 1930s (see Feest 2005). Based on these historical case studies I have argued that even though the psychologists in question saw themselves as closely allied with logical positivists, they in fact pursued a rather different intellectual project. Their research was not aimed at providing concepts with meaning, but to conduct empirical investigations of the referents

of concepts. In order to engage in this dynamic project, they had to temporarily fixate conditions of application for the concepts in a way that would allow them to perform experimental interventions.

In the recent literature in the philosophy of experimental neuroscience, my account has found some support in an article by Jacqueline Sullivan (2009), who shares my contention that “[a]n operational definition is built directly into the design of an *experimental paradigm*” (Sullivan 2009, 514). I take this way of talking about experimental paradigms to be highly congenial to my own formulation (see above) that operational definitions provide *paradigmatic conditions of application* for a given concept, thereby specifying a standard procedure for the scientific investigation of the phenomenon thought to be picked out by the concept (as distinct from the more specific details of how this is implemented in particular experiments).³ For example, as we will see in Section 5 below, a common experimental paradigm for the study of memory calls for a distinction between a study phase and a test phase. This paradigm presupposes a particular understanding of memory as a phenomenon whereby previously learned material can be retained and retrieved. Such an underlying preconception is encapsulated in an operational definition, which might state that if a subject is exposed to a particular learning task then memory is present if and only if retention of the learned material can be demonstrated in the test phase. My question in this paper is how the fact that an operational definition is built into the design of an experimental paradigm makes it possible to generate knowledge *by means of* that experimental paradigm. My thesis is that operationally defined concepts provide scientists with instructions about experimental interventions designed to produce data that indicate the phenomenon of interest. This is analogous to the way in which a physical tool might come with instructions on how to physically intervene in the world in order to achieve a particular goal.

Now, one might raise the question of whether the definition in question is a *valid* tool. In other words: does the fact that a given operational definition specifies empirical conditions of application for a given concept mean that the concept in fact picks out the phenomenon of interest? The answer is a very clear “no”: “[t]he strategy of investigators ... is to *assume* that the operationalizations that they provide are actually indicative of the function of interest” (Sullivan 2009, 518). This assumption may, of course, turn out to be false, just like the assumption that a particular screwdriver is adequate to the purpose of assembling a given piece of furniture may turn

³ See Sullivan (2009) for a more detailed analysis of the distinctions between experimental designs, paradigms, protocols and sub-protocols.

out to be false. While questions of validity are clearly pertinent, they are not addressed in this paper. Here my focus is not on the question of how the usage of a concept can be justified, but on providing a descriptively accurate account of the practical work that is done by operationally defined concepts in specific experimental contexts.

IV. ON THE NOTION OF A SCIENTIFIC TOOL

Up to this point I have used the word “tool” rather loosely. It will probably be objected that the analogy between concepts and tools or instruments is not very compelling, since (a) concepts are not physical devices like hammers or screwdrivers, and (b) hammers and screwdrivers do not perform the kind of epistemic function I have attributed to concepts, and (c) when we use a physical tool, such as a screwdriver, we typically know what kind of object it can be used on (e.g., a screw), but what are we to make of a tool for “ill-understood” objects? In this section I hope to further clarify my account by responding to all three objections. Answering to the first charge, I will argue that there is no good reason to limit the meaning of the word “tool” to physical devices. Answering to the second charge, I will argue that in the sciences there are in fact plenty of tools that do perform epistemic functions. Finally, in response to the third question, I will explain my notion of the relationship between tool and object by comparing it with Rheinberger’s notion of an *epistemic object*.

Why Tools Do Not Have to be Physical

A tool is a device that enables us to do something. In the context of empirical research, we are interested in tools that enable scientists to generate and analyze data. The latter of these two functions, that is, that of analyzing data, already makes it quite clear that scientific tools do not have to be physical devices, though they may make use of physical devices in their execution. For example, when a scientist uses a statistical rule of inference to test whether a particular empirical effect is significant, it is fair to refer to this method as a tool, even though it is not a physical machine. This does not necessarily mean that there are no physical machines involved, as the rule of inference is often implemented in a physical device, for example, a calculator or computer, which delivers the results of the data analysis.

In the present context my focus is not on data-*analyzing* tools, but on data-*generating* tools. My claim is that in psychology there are data-generating tools that are not physical devices, at least not of any very technologically sophisticated kind. A typical example of such tools are psychological tests or questionnaires. There is, of course, a material

realization of such tools, insofar as the test is presented on a piece of paper or a computer screen, and insofar as the subjects' responses are recorded in some way or other. But the epistemic work of the tool is done by the ways in which the questions are asked, not by the physical medium in which they are presented. In this respect I am in full agreement with Sturm and Ash (2005), who write that "there is no good reason to define the concept of instrument or of psychological instruments in such a way that only physical devices, let alone technologically advanced ones, are included" (Sturm and Ash 2005, 15). But even if this is true, it does not follow that *concepts* can be instances of such non-physical and non-mechanical tools. Nor does it follow that they are tools for knowledge generation. In other words, my argument requires two more steps. First, I need to specify the kind of tool that I wish to compare concepts with. Second, I need to say more about the notion of knowledge-generation presupposed here.

Tools that Perform an Epistemic Function

Some tools can provide us with knowledge. Another way of putting this is to say that in the sciences tools are often not simply regarded as generating data, but as generating knowledge by means of data. The question is what has to be in place such that they can do this. Let us approach this question by looking at two groups of instruments of which this is obviously true, namely measuring instruments and instruments designed to enhance our observational capacities. Examples of the first kind are thermometers, scales, and stop watches. Examples of the second kind are microscopes or telescopes. Such tools can provide us with knowledge about the amount or degree to which a given quality (e.g., temperature or weight) is present, or about the presence of entities or phenomena that wouldn't be detectable by means of the "naked" human senses alone.⁴

But is there a sense in which a similar claim can be made about concepts? I argue that there is. This is especially clear in psychology, where, as mentioned above, common tools are psychometric *tests*, such as intelligence tests. When we think of the notion of *intelligence*, features that come to mind are logical reasoning, analytical skills, and creative problem solving, each of which are exemplified by certain types of behaviors in particular circumstances. Given our knowledge of such typical indicators of intelligence, it is not far-fetched to try to devise a standardized test that measures intelligence by way of those indicators, thereby allowing us to determine the degree of intelligence of a particular person. In other

⁴ Heidelberger (1998) refers to these two types of instruments as representational and productive tools, respectively.

words, what are regarded as typical behavioral indicators of a trait like intelligence are going to make their way into a test that is designed to measure this trait. The presuppositions in question are conceptual. And the test may be regarded as a psychological tool. This suggests a close relationship between the two, in that psychological tools like intelligence tests are implementations of particular types of conceptual assumptions. For example, say that our notion of intelligence involves the operational definition that if a person X is presented with a certain type of task they will be regarded as intelligent if and only if they do well on that task. This corresponds directly to a possible tool for determining a person's intelligence.

What was just said will immediately provoke the response that this kind of operationism or logical behaviorism not only relies on a dated theory of meaning, but also defines the property under investigation by means of the test, thereby making it hard to see how the test could help us find out something new about the property in question. In response, I would like to emphasize again that I am not claiming that the meaning of a concept like intelligence can be reduced to the response behavior on a psychometric test. If it did, then, trivially, a test would not have to be validated, since the test results, by definition, would instantiate the property measured by the test. As mentioned earlier, my focus here will not be on the issue of validation. Let me point out, however, that already in the methodological discussions of the 1950s, some psychologists acknowledged that a simple operational or criterion-oriented validation was not going to be sufficient, since the objects or phenomena of interest could usually not be exhaustively delineated by an empirical criterion. Figuring out what the test was in fact measuring, therefore, required some additional theoretical work, which the authors referred to as "construct validation" (Cronbach and Meehl 1955).

With respect to the question of how to make sense of the epistemic function of tools, given the fact that some such tools already makes conceptual assumptions about their objects, we need to distinguish between two types of epistemic interests (and, consequently, two types of epistemic functions a tool might play): (1) that of finding out whether, and to what extent, the phenomenon X is present in a given circumstance, and (2) that of exploring the very nature of a given phenomenon X. If our interest is in the first type of question, we will typically use an instrument that can detect instances of the phenomenon, where this instrument will embody some preconception about the phenomenon. Justifying that preconception is what the validation of the instrument is all about. If, however, our interest is in the second type of question, things are more complicated, since here we have to rely on a tool prior to having validated it. It is this latter type of

case that the thesis of this article is about.

How Tools Can Be Applied to Ill-Understood Objects

Some tools can provide us with knowledge about ill-understood objects. In the previous subsection I have argued that tools can have an epistemic function, and that concepts can be such tools, illustrating both points with the example of psychometric tests. I concluded by distinguishing between two types of epistemic functions tools might play: that of measuring or detecting instances of phenomena, given a tool that has already been validated or standardized, and that of empirically exploring the nature of the phenomena prior to having a validated and/or standardized tool. I drew attention to the fact that in the latter kind of case scientists may only have a vague or inaccurate conception of the phenomenon they are trying to investigate. My thesis is that such a vague or inaccurate conception of the phenomenon under investigation can be epistemically productive, provided that it is cast in the form of an operational definition that allows researchers to intervene in the domain of study.

Let me conclude this section by briefly explaining my notion of a “purported” or “ill-understood” object or phenomenon. This notion may appear odd since it seems to suggest that scientists can investigate objects whose very existence is in question, that is, that they could conceivably investigate non-existent objects. Clearly, this is not what I have in mind. What I have in mind, rather, is that scientists can take themselves to be investigating objects, which—on closer inspection—turn out not to exist, or to be very different from the way they were thought to be. What this suggests is that we need to distinguish between two notions of “object”: (1) objects as mind-independent things in the world, and (2) objects of research *as conceptualized by scientists*. Scientific conceptualizations of objects or phenomena in the world can, of course, be inaccurate, incomplete, or vague. Likewise, mind-independent objects may not be very well understood. This draws attention to the fact that scientists typically think of the objects and phenomena of their research as mind-independent, yet the only way they have access to those objects and phenomena is through their own conceptualizations. The question is how such conceptualizations can be productive in the generation of knowledge about what scientists will typically consider mind-independent objects.

The approach taken in this article bears some affinity with Hans-Jörg Rheinberger’s account, in that he, too, focuses on the dynamics of knowledge generation about specific objects (Rheinberger 1997). In addition, my notions of purported or ill-understood objects are reminiscent of Rheinberger’s *epistemic objects*, which he defines as objects that attract

our epistemic curiosity. I am also sympathetic with Rheinberger's idea that an epistemic object can be turned into a technical object once it is no longer investigated for its own sake, but rather used as a *tool*. As he emphasizes, a technical object can, at any point in time, go back to being an epistemic object. This can happen when the functioning of the tool itself attracts our epistemic curiosity rather than being taken for granted. Where my approach parts ways with Rheinberger's however, is that he does not distinguish between objects and *concepts* of objects. This means that it is impossible, within his approach, to capture my contention that concepts can be tools for the study of objects.⁵

V. THE CONCEPT OF IMPLICIT MEMORY AS A TOOL

In this section I will turn to a concrete example to illustrate my thesis about the role of concepts as tools. The research I want to introduce here has as its purported object the phenomenon of *implicit memory*. The expression "implicit memory" was introduced into the literature in 1985 by the psychologists Graf and Schacter, following the discovery of a particular *data pattern*, which resulted from a particular set of *experimental manipulations*. The relationship between the experimental manipulation and the data pattern subsequently provided researchers with an operational definition of a concept. It is with this implicit conceptual assumption that they proceeded to investigate the purported phenomenon (implicit memory). In order to understand the nature of the research it is important first to explicate the rationale of experimental research on human memory.

Experiments in memory research typically consist of three parts: (1) the study phase, (2) an intermediate phase and (3) the test phase (see Lockhart 2000). During the *study phase*, subjects are exposed to some experimental items (e.g., words, pictures, nonsense syllables) so as to enable them to memorize them. The conditions under which this takes place can be varied. For example, subjects might simply be asked to memorize the items. Or they might be instructed to use particular strategies, or they might be given some task without being told that the aim is to memorize the items before them. The actions that researchers take in order to create such learning conditions are also referred to as "experimental manipulations." During the *test phase*, memories of the experimental items are elicited from the subject by means of a memory test. There are different kinds of memory tests. The kinds of memory tests that have traditionally been used are either recall or recognition tests.

⁵ For a more detailed discussion of the relationship between my approach and Rheinberger's program of historical epistemology see Feest, under review.

When subjects take such tests, they are *explicitly* asked to report what they remember—or recognize—from the study phase. For this reason this type of test is called an “explicit test.” Such tests can be contrasted with another type of test, which has been used more recently (within the last forty years or so). An example of such a test is one where the subject has to decide whether a given string of letters on the computer screen is a word. The subject is judged as having a memory of an item if the reaction time is shorter for words from the test phase than for others. The subject is not explicitly instructed to recall something. For this reason this type of test is referred to as an “implicit test.”

It has been known since the early 1970s that prior exposure to an experimental item increased the likelihood of doing well on an implicit test for that item, even if the subject has no explicit recollection of the exposure. In the early 1980s, there was increasing evidence for the existence of so-called “experimental dissociations” between the results of explicit and implicit memory tests. This means that the results on the two kinds of tests differed *even though the learning conditions in the study phase had been identical*. For example, (1) a normal subject has difficulties explicitly remembering verbal items a week after the learning phase, while performance on an implicit memory test is as good as it was right after the learning phase,⁶ and (2) results on explicit tests can be influenced by certain variations of the conditions during the study phase (for example, elaborative study conditions), which do not appear to affect the performance on implicit tests. (3) The experimental dissociations were found in patients that suffered from amnesia (typically patients with lesions in the hippocampus).

The discovery of the dissociation gave rise to experimental studies of the phenomenon, which soon came to be called “implicit memory.” My thesis here is that the conditions that produced the dissociation were treated as an operational definition of the concept of the phenomenon, and that this definition, in turn, functioned as an (evolving) tool for the study of the phenomenon. The first of the above conditions represents the most basic definition:

$$Q_1 \rightarrow (Q_3 \text{ iff } Q_2) \quad (\text{modeled on Carnap 1936})$$

(If a human subject is exposed to a given learning paradigm and then tested with an explicit and implicit memory test, then implicit memory is present iff there is a functional dissociation between the results on the implicit and the explicit memory test.)

⁶ “Normal,” in the context of psychological research, usually means “not brain-damaged.”

When we look at how it was in fact used in the further research about the purported phenomenon, it is clear that the concept was further fine-tuned in terms of the other two conditions (see above) and this in turn played a role in research. This can be illustrated by means of studies that were reported by Graf and Schacter (1985). The authors reported dissociations between implicit and explicit memory tests. However, on the other hand they also reported that the priming effect was more pronounced under the elaborative study condition than under the non-elaborative study condition. This went against the assumption that implicit memory is unaffected by variations in the study condition. Consequently, the authors discussed the possibility that the observed dissociation *might not* be due to implicit memory. In order to investigate this question, they then conducted another experiment in which they replicated the experiment with patients who suffered from amnesia. This experiment once again produced the experimental dissociations. As a result of that experiment, the authors then cautiously concluded that their findings were indicative of implicit memory after all, thereby obviously relaxing the assumption that implicit memory could not be affected by study conditions.

This case is illuminating in several respects. First, the example illustrates that despite the rather fragmented understanding of the phenomenon under investigation, there was early on a presupposition to the effect that there was indeed *one* phenomenon to be studied, such that it could be asked whether a given experimental result was indeed indicative of the phenomenon. Hence, we may say that there was a general understanding amongst researchers that they were all using the expression “implicit memory” to refer to one and the same thing. Second, it illustrates the function played by paradigmatic conditions of application of the expression in the early stages of investigating the corresponding (purported) phenomenon. Hence, we may say that there was a shared understanding that the phenomenon under investigation was to be empirically individuated by means of particular experimental operations, insofar as they resulted in particular data. Third, it shows that researchers were willing to relax one of their initial assumptions, thereby revising the initial concept. Hence, we may say that at least some presuppositions of the research were changed as a result of the research.

VI. THE TOOL IN CONTEXT

Having looked at the way in which the *typical criteria of application* may be modified or adjusted in the research process, I would like to conclude with a brief discussion of the question of whether and how *ideas about the referent* underwent changes as research on the phenomenon proceeded. This is to show two things: First, the fact that no full-blown theory of

implicit memory existed when the research started, and that scientists had to revert to operational definitions in order to empirically individuate instances of the purported phenomenon, does not mean that we can neglect the role of theoretical considerations in this research. Second, the way in which theoretical ideas about the nature of the referent changed illustrates the kind of scenario already mentioned above, namely that there is no a priori way of predicting the effects of changing ideas about the nature of a phenomenon, that is, whether scientists will conclude that the phenomenon they thought they were investigating does not exist, or whether they will change their minds about what the phenomenon is like.

When the dissociation between performances on explicit and implicit memory tests was first discovered, memory was discussed in terms of long- and short-term memory (e.g., Baddeley 1976). Within neuropsychology, it had also been noted that amnesiacs appeared to exhibit dissociations between long- and short-term memory performances (Baddeley and Warrington 1972). The distinction between long- and short-term memory was supplemented with the notion that there might be two kinds of long-term-memory, one “episodic” and one “semantic” (e.g., Tulving 1983), where the former was thought to be responsible for our ability to remember biographical episodes, whereas *semantic memory* was thought to be memory for facts, independently of how knowledge of such facts was acquired. The items in semantic memory were thought to be stored in a network of interconnected nodes, according to their semantic relationships. Faced with the experimental dissociations between reports on explicit and implicit tests, reported above, Tulving et al. (1982) hypothesized that performance on recognition tests (i.e., on *explicit* tests) was supported by a recollection of the study phase, that is, by episodic memory. This raised the question of whether performance on the priming test (i.e., on *implicit* tests) might be due to semantic memory. However, by 1982, it was no longer clear whether this explanation could hold up, as there was now also evidence for *sensory* priming (i.e., that the priming effect occurred *not* by virtue of the meaning of a word, but by virtue of some more superficial feature, such as the way it was being presented; e.g. Jacoby and Dallas 1981). This led some to think of implicit memory as such a memory *system*, which was responsible for both semantic and sensory priming.

Subsequent researchers in the field can roughly be divided into two groups: One group questions whether the data in question (i.e., the priming data) should be viewed as indicative of *memory* phenomenon at all (e.g., Roediger 2003). The other group of theorists still talks about implicit memory systems (e.g., Schacter 1999). However, Schacter now thinks of implicit memory as a perceptual representation system (PRS),

a system that has several domain-specific subsystems, which are each responsible for processing different sensory features of the experimental items (thereby explaining the dissociations between and within sense modalities), all of which work independently of the semantic (explicit) processing of the items (thereby explaining the dissociations between explicit and implicit tests).

The point here is that even among those who believe in the existence of an implicit memory system, the kind of phenomenon they now believe exists is significantly different from what they believed implicit memory was twenty-five years ago. While the term was initially supposed to refer to a system, which was taken to function independently of whatever systems process specific modalities, it is now (by some) taken to refer to a system that specifically processes *sensory* aspects of experimental items. Thus, the scientific conception of this phenomenon appears to have incorporated some of the features that it was initially contrasted with. It bears stressing, though, that the changed conception of the referent was brought about, in part, by research that relied on a particular operational definition, that is, on a particular preconception about how to empirically identify the purported referent by means of an experimental operation. I have argued that this (preliminary) preconception functioned as a research *tool*.

VII. SUMMARY

I would like to conclude by placing the topic of this paper in the broader context of some issues in philosophy of experimentation. Starting with Ian Hacking in the early 1980s, the line of debate has often been cast as being between those who emphasize the primacy of theory and those that emphasize that experiments have a life of their own (Hacking 1983). Hacking in particular took the shift from theory to experimental practice to also be a shift away from a philosophical preoccupation with language. While sharing some of Hacking's fundamental intuitions, my aim in this article has been to put language back on the agenda. Even if we don't think that the primary aim of experiments is theory-testing, we still have to make room for the way in which scientific conceptualizations enter into the experimental process. On the thesis argued for here, preliminary scientific concepts may be viewed as tentative representations of purported objects. However, more importantly, they are tools that allow for the experimental *intervention* into the domain of study. As I showed, such tools may well be informed by existing theories of the domain. Yet, they also function in a space that is more open and dynamic than suggested by traditional notions of theory and theory-testing. By providing an analysis of this open and dynamic quality in terms of the notion of operational definitions, I have shown how concepts can figure as tools for the experimental generation of

knowledge.

ACKNOWLEDGMENTS

An earlier version of this paper was presented at the conference “Generating Experimental Knowledge” in Wuppertal, Germany (June 2007) and at the 2008 meeting of the PSA in Pittsburgh (November 2008). The author would like to thank members of both audiences for insightful comments and discussions. Many thanks go also to three anonymous reviewers for this journal. The paper has been revised to incorporate their helpful suggestions.

ULJANA FEEST
Technische Universität Berlin
feest@mail.tu-berlin.de

REFERENCES

- Andersen, Hanne, Peter Barker, and Xiang Chen. 2006. *The Cognitive Structure of Scientific Revolutions*. Cambridge: Cambridge University Press.
- Baddeley, Alan. 1976. *The Psychology of Memory*. New York, NY: Harper & Row.
- Baddeley, Alan, and Elizabeth Warrington. 1970. Amnesia and the Distinction Between Long- and Short-Term Memory. *Journal of Verbal Learning and Verbal Behavior* 9: 176-89.
- Bridgman, Percy. 1927. *The Logic of Modern Physics*. New York: Macmillan.
- Bogen, James, and James Woodward. 1988. Saving the Phenomena. *Philosophical Review* XCVII, no. 3: 303-52.
- Carnap, Rudolf. 1936. Testability and Meaning. *Philosophy of Science* 3, no. 4: 419-71.
- Chalmers, David. 1996. *The Conscious Mind: In Search of a Fundamental Theory*. New York : Oxford University Press.
- Chang, Hasok. 2009. Operationalism. *The Stanford Encyclopedia of Philosophy* (Fall 2009 Edition), ed. Edward N. Zalta. plato.stanford.edu/archives/fall2009/entries/operationalism.
- Cronbach, Lee, and Paul Meehl. 1955. Construct Validity in Psychological Tests. *Psychological Bulletin* 52, no. 4: 281-302.
- Ebbs, Gary. 2003. Denotation and Discovery. In *Socializing Metaphysics: The Nature of Social Reality*, ed. Frederick F. Schmitt, 247-68. Lanham: Rowman and Littlefield.
- Feest, Uljana. 2005. Operationism in Psychology—What the Debate is About, What the Debate Should Be About. *Journal for the History of the Behavioral Sciences* XLI, no. 2: 131-50.
- Feest, Uljana (under review). Remembering Memory: The Death of an Epistemic Object?

- Graf, Peter, and Daniel Schacter. 1985. Implicit and Explicit Memory for New Associations in Normal and Amnesic Subjects. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 11, no. 3: 501-18.
- Hacking, Ian. 1983. *Representing and Intervening. Introductory Topics in Philosophy of Science*. Cambridge: Cambridge University Press.
- Heidelberger, Michael. 1998. Die Erweiterung der Wirklichkeit im Experiment [The extension of reality in experimentation]. In *Experimental Essays–Versuche zum Experiment*, eds. M. Heidelberger and F. Steinle, 71-92. Baden-Baden, Germany: Nomos Verlagsgesellschaft.
- Jackson, Frank. 1998. *From Metaphysics to Ethics: A Defense of Conceptual Analysis*. Oxford: Clarendon Press.
- Jacoby, Larry, and Mark Dallas. 1981. On the Relationship Between Autobiographical Memory and Perceptual Learning. *Journal of Experimental Psychology, General* 110: 306-40.
- Laurence, Stephen, and Eric Margolis. 1999. Concepts and Cognitive Science. In *Concepts: Core Readings*, eds. Eric Margolis and Stephen Laurence, 3-81. Cambridge, MA: MIT Press.
- Lockhart, Robert S. 2000. Methods of Memory Research. In *The Oxford Handbook of Memory*, eds. Endel Tulving and Fergus I.M. Craik, 45-57. Oxford: Oxford University Press.
- Putnam, Hilary. 1975. The Meaning of 'Meaning.' *Mind, Language and Reality: Philosophical Papers, Vol. 2*, 215-71. Cambridge: Cambridge University Press.
- Rheinberger, Hans-Jörg. 1997. *Towards a History of Epistemic Things: Synthesizing Proteins in the Test Tube*. Stanford: Stanford University Press.
- Roediger, Henry L. 2003. Reconsidering implicit memory. In *Rethinking Implicit Memory*, eds. Jeffrey S. Bowers, and Chad S. Marsolek, 3-18. Oxford: Oxford University Press.
- Rosch, Eleanor. 1978. Principles of Categorization. In *Cognition and Categorization*, eds. Eleanor Rosch and Barbara B. Lloyd, 27-48. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Schacter, Daniel. 1999. Implicit vs. Explicit Memory. In *The MIT Encyclopedia of Cognitive Sciences*, eds. Robert Wilson and Frank Keil, 394-95. Cambridge, MA: MIT Press.
- Sullivan, Jacqueline. 2009. The multiplicity of experimental protocols: a challenge to reductionist and non-reductionist models of the unity of neuroscience. *Synthese* 176: 511-39.
- Sturm, Thomas and Mitchel Ash. 2005. Roles of Instruments in Psychological Research. *History of Psychology* 8, no. 1: 3-34.
- Tulving, Endel. 1983. *Elements of Episodic Memory*. New York: Oxford University Press.
- Tulving, Endel, Daniel Schacter, and Heather A. Starck. 1982. Priming Effects in Word-Fragment Completion are Independent of Recognition Memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition* 8: 336-42.