Observational and Theoretical Terms in Psychology

A Cognitive Perspective on Scientific Language

James M. Clark and Allan Paivio
University of Western Ontario

ABSTRACT: This article presents an empirical, psychological approach to the distinction between observational and theoretical terms, a persistent and unresolved issue in the study of science. Empiricists maintain that observational terms are more directly connected to perceptual experiences than theoretical terms and have more reliable and definite meanings. A contrasting rationalist view is that all scientific terms are equally theory laden and unreliable. Most studies of this problem have been philosophical, but the problem is also amenable to empirical investigation. Using relevant theory and evidence from cognitive psychology, we show that the observational–theoretical distinction parallels the concrete–abstract distinction in natural language, that both follow from the dual coding assumption of distinct verbal and nonverbal cognitive systems, and that natural language research supports this view. We report similar evidence for 72 psychological terms that were rated by psychologists on observability, consistency of meaning, and other relevant attributes. The ratings correlated highly with one another and also with imagery reaction times and reports of spontaneous imagery. The importance of observability in science is further supported by the relative concreteness of scientific communication and related phenomena. The evidence challenges the claim that all scientific terms are equally inferential and reliable and is consistent with empiricist views of science that emphasize the distinct properties of concrete, observational terms and abstract, theoretical terms.

The distinction between observational and theoretical scientific terms is a controversial issue that appears explicitly or implicitly in many writings about science. We survey the conceptual background of the observational–theoretical distinction as presented in the traditional literature. Although stimulated by thoughtful philosophical arguments, the debate remains unresolved, perhaps because some of the contentious issues involve complex empirical and psychological claims about the language and cognition of scientists. To understand such issues about science, "Psychology has no one to look up to. It must turn to itself, it being the eye that looks into itself, to see how it sees" (Wolman, 1971, p. 885). Accordingly, we examine cognitive theory and evidence that is relevant to the distinction between observational and theoretical terms. This psychological approach sheds new light on the debate and could ultimately reveal the proper place of the distinction in the understanding and practice of science.

The observational–theoretical debate goes beyond claims that theory determines what scientists choose to observe and goes to the very heart of cognitive aspects of science. Psychologists use many terms: *test score* and *intelligence, heart rate* and *anxiety, proximity* and *attitude, eye movement* and *attention,* and so on. As we will show, such psychological terms, like terms in other sciences, vary in how closely tied they are to concrete observations. Proponents of the observational–theoretical distinction believe that the distinction is valid and properly plays an important role in scientific thinking and communication, and they attach greater certainty and intersubjective agreement to observational terms (e.g., test score and heart rate) than to theoretical terms (e.g., intelligence and anxiety). Others minimize the importance of the distinction and maintain that observational and theoretical terms are equally inferential and do not warrant different degrees of confidence. At its most fundamental level, then, the debate concerns whether there is a concrete level of observation and associated discourse about which scholars can agree despite their differing theoretical orientations. Contrasting answers to this basic question can lead to radically different views of science and recommendations for its practice.

Because the issues are complex, it may facilitate understanding to lay some foundations at this point. First, the theory and research presented in this article focus on the defining and correlated properties of observational and theoretical terms and leave unanswered many complex questions raised in philosophical discussions of the observational–theoretical issue. This emphasis reflects our belief that basic aspects of observational and theoretical terms are fundamental to empirical investigation and understanding of more complex aspects of scientific language. Second, our primary concern is with cognitive aspects of scientific language; our work may not address unique concerns of philosophers that differ from those of psychologists (see Kendler, 1981, pp. 5–6). Third, claims about distinct properties of observational and theoretical terms derive in part from putative differences between the observations and theoretical constructs to which such terms refer. For that reason, we sometimes appeal to relevant nonlanguage aspects of science. Fourth, we emphasize psychological language, especially in the empirical section of the article, but subsequently show that other sciences demonstrate phenomena consistent with our conclusions.

 0003-066X/89/$00.75
Vol. 44, No. 3, 500-512

Background in Philosophy of Science

Some students of science give a central role to observational statements that describe experience, claiming that such statements provide a more objective foundation for knowledge than do theoretical statements (e.g., Hempel, 1952; Nagel, 1961; Quine & Ullian, 1978). The distinction between observational and theoretical terms was a central feature of classical empiricism (Sklar, 1985; Suppe, 1977) and remains important in contemporary discussions of science (e.g., Churchland & Hooker, 1985). This received view of science had a profound influence on psychology (see Flugel, 1964) and other sciences. In psychology, the observational–theoretical distinction contributed to operationism ("Symposium on Operationism," 1945), construct validity (Cronbach & Meehl, 1955), and other basic methodological developments. Psychology became distinguished from nonscientific approaches to the study of mind by the priority given to empirical evidence that is objective, based on the senses, and observational (e.g., Hull, 1943; Kerlinger, 1979; Selltiz, Wrightsman, & Cook, 1976; Underwood, 1957).

After psychology had explicitly adopted the distinction between observational and theoretical terms, contrasting views became dominant in the philosophy of science. Kuhn (1962) and other scholars (see Landesman, 1970; Morick, 1980) rejected the superiority attributed to observation statements and maintained that such statements were not privileged, independently credible, free from error, or noninferential. Observational statements were no longer viewed as the objective, true, or reliable foundation for theory (Brown, 1977; Feyerabend, 1969), and the distinction between observational and theoretical terms lost its traditional value (Kuhn, 1979). The arguments led various psychologists similarly to question or even abandon views of science that assign a special status to observation (e.g., Borgen, 1984; Gergen, 1980; Rorer & Widiger, 1983; Segal & Lachman, 1972). The strongest rejections stated, perhaps rhetorically, that empiricism is "pathological" (Koch, 1981) and an "atheoretical cul-de-sac" (Royce, 1982, p. 264). Others argued that theory precedes observation (e.g., Meyers & Grossen, 1974) and that data are not supreme (Mahoney, 1976), or proposed top-down approaches to science that diminished the role of observation (e.g., Pylyshyn, 1979). Several psychologists have reviewed these debates (e.g., Kendler, 1981; Weimer, 1979).

Despite objections to the observational–theoretical distinction, the controversy has never been completely settled. Not all scholars of science rejected the distinction (e.g., Losee, 1980; van Fraassen, 1980, 1985), and there are indications of a renewed appreciation of the central role of observation in science (e.g., Fodor, 1984; Sklar, 1985). Many psychologists were also unconvinced about the alleged shortcomings of observation as a foundation for science (e.g., Broadbent, 1973; Kendler, 1981). Cook and Campbell (1979), for example, asserted that observation had become almost irrelevant in contemporary philosophy of science and suggested instead that observation be emphasized more than theory until dependable facts were available. Meehl (1978) likewise questioned the dismissal of objective views of science.

This unresolved debate has been largely philosophical in nature and based on rational arguments, but research and theory in cognitive psychology are also relevant to the controversy. An empirical approach is justified because (a) the disputed differences between observational and theoretical terms are relative rather than absolute, and (b) the common assumption that the observational–theoretical distinction requires only philosophical rather than empirical justification appears to be incorrect. These claims are considered in turn.

Absolute Versus Relative Views of the Distinction

The major defining attribute of the observational–theoretical distinction was the observability of the entities and events to which scientific terms refer. Proponents of the distinction claimed that terms vary in the directness of their connections to perceptual referents. That is, such terms as *heart rate* and *meter reading* were alleged to refer to directly observable things or traits, and such terms as *schema* and *field* to hypothetical entities that are neither directly observable nor measurable by simple procedures (Carnap, 1966; Nagel, 1961; Suppe, 1977; Toulmin, 1953). Similar grounds have been used to differentiate observable psychological concepts from abstract, theoretical concepts (e.g., Anderson, 1971; Carlsmith, Ellsworth, & Aronson, 1976).

Opponents of the distinction responded that both observational and theoretical terms are known only indirectly (Popper, 1965) and are equally theory laden (discussed by Cook & Campbell, 1979; Hanson, 1958). According to this view, inference plays the same role in the definition of terms for physical objects (e.g., *tree*) as in the definition of such theoretical terms as *electron* (Bronowski, 1978; Quine, 1951). Instances of perception in which subjective factors operate (e.g., ambiguous stimuli perceived as either rabbits or ducks) were presented as paradigm cases of theory-laden percepts and as evidence against the observational–theoretical distinction (e.g., Brown, 1977; Morick, 1980). Because the meanings of all scientific terms were thought to be indirect and theory laden, observability was considered an inappropriate way to classify scientific terms. Consistent with this conclusion, it was demonstrated that the proposed criteria for the distinction, notably observability, did not permit unambiguous classification of scientific terms (see Achinstein, 1968, ch. 5; Suppe, 1977, pp. 66–86).

The case against the observational–theoretical dis-

This work was supported by Grant No. A0087 from the Natural Sciences and Engineering Research Council of Canada to A. Paivio. We appreciate comments on earlier drafts of the article by Ian Begg, Trudy Bons, Bill Harper, Robert Hoffman, Albert Katz, Ed Manukian, Jim Neufeld, Mary Walsh, and several anonymous reviewers. We also thank Trudy Bons and Ed Manukian for their fine work on this project, and the many psychologists who provided data for our studies.

Correspondence should be addressed to James M. Clark, University of Western Ontario, Faculty of Social Science, Department of Psychology, London, Ontario, Canada, N6A 5C2.

tinction has been challenged on several grounds (e.g., Kendler, 1981; Martin, 1972). Fodor (1984), for example, has pointed out that perceptual processing can be remarkably resistant to penetration by beliefs and theories. Some of the strongest arguments against the distinction are problematic only for claims that observational terms are absolutely objective and that the distinction between observational and theoretical terms is completely unambiguous. The arguments carry less weight against an alternative view that observational and theoretical terms differ relatively rather than absolutely. For example, although ambiguous stimuli challenge the extreme claim that all observations are completely objective and noninferential, such cases do not warrant the conclusion that "observational terms are just as theory-laden as any others" (Churchland, 1985, p. 36).

Major proponents of the observational–theoretical distinction explicitly expressed a relative view. They stated that no sharp boundary or precise criterion separates the two classes and that observable and nonobservable were points on a continuum (Carnap, 1966; Nagel, 1961). That theoretical terms may ultimately become part of the observational language (Carnap, 1966; Shapere, 1985) further blurs the boundary between observational and theoretical terms. However vague and uncertain, the distinction between observational and theoretical terms was nevertheless thought to be valid and to have implications for the understanding of science (Nagel, 1961, p. 83; Newton-Smith, 1981; van Fraassen, 1980). This belief is consistent with the philosophical view that such fuzzy distinctions are meaningful (e.g., Michalos, 1970; Thouless, 1974) and with the contemporary cognitive views that many concepts are represented by prototypes or probabilistic features rather than by necessary and sufficient absolute criteria (Smith & Medin, 1981). "Family resemblances" may be especially relevant to such abstract, philosophical categories as *observational* and *theoretical.*

Thus, fuzzy views of the observational–theoretical distinction cannot be decided by examples of subjectivity in perception that are inconsistent with absolute objectivity. Relative claims require empirical data that would determine whether scientific terms vary reliably on observability and whether observability correlates, albeit imperfectly, with such relevant properties as intersubjective agreement. However, this empirical approach to the observation–theory issue has itself been questioned.

Foundations of Science and the Problem of Circularity

Much of the traditional literature on the observational–theoretical distinction assumes that scientific activities are to be justified by reason and not by empirical study. It has been argued that observation presupposes empiricism and that empirical studies of science and its language therefore involve a vicious circle. Gergen (1986), for example, asserted that empirical evidence cannot prove or falsify claims about facts because "using the paradigm to justify itself ... [rules] out the possibility of falsification" (p. 481). In a similar vein, Messer (1985) challenged empirical investigations of psychologists (e.g., Kimble, 1984; Krasner & Houts, 1984) on the grounds that empirical methods predetermined their outcomes. Beliefs that science cannot study itself may also explain why the traditional literature on the observational–theoretical distinction is largely nonempirical and why the metatheoretical foundations of psychology and other sciences are usually sought in philosophy rather than in scientific psychology itself. Although the circularity argument seems compelling, a contrasting view is that empirical methods can and should be used to study science (for example, Campbell, 1977, cited in Brewer & Collins, 1981; Mahoney, 1976; Mandler & Kessen, 1959; Stevens, 1939; Tweney, Doherty, & Mynatt, 1981; Weimer, 1979; Wolman, 1971). A priori, empirical studies of science seem at least as justified as rational studies, as shown in the next few paragraphs.

The basic claim that empirical methods necessarily determine their own outcomes is not compelling for several reasons. Most important, the empirical evidence for this putative circularity is lacking, and it is not obvious what findings would justify such a strong conclusion. It is also difficult to understand how empirical studies of science necessarily predetermine their results. For example, we later report high correlations between rated observability and rated consistency of meaning of psychological terms. Nothing prevented the relevant correlations from being zero or even negative, which would falsify predictions about science derived from the observational–theoretical hypothesis. Even granting that empirical results are sometimes affected by experimenter beliefs (Rosenthal, 1966) or other confounding variables, such competing explanations are not limitations in principle of empirical studies of science and are themselves amenable to and only revealed by experimental testing and control. Finally, if circularity per se were a valid argument against using scientific methods to investigate science, then rational approaches to the study of science would be similarly tainted, especially because reason and intuition seem even more susceptible than observation to the influence of expectations.

An empirical psychological approach to the observational–theoretical distinction can be justified further by evidence that systematic observations overcome relevant limitations of reason (e.g., Kahneman & Tversky, 1973; Michalos, 1970; Thouless, 1974). For example, people find it difficult to discover imperfect covariation between variables under optimal conditions (Nisbett & Ross, 1980; Peterson & Beach, 1967), and complex situations aggravate the problem. But the observational–theoretical issue specifically concerns imperfect relations between such constructs as observability, theory-ladenness, and intersubjective agreement. Moreover, these relations are obscured by other complex phenomena of science. To evaluate relative claims about imperfect covariation, one must operationally define constructs, obtain systematic observations, and correlate the critical variables. That these empirical methods are brought to bear on science itself does not disqualify them.

We conclude that empirical study of the observa-

tional–theoretical hypothesis can make unique contributions to the understanding of science and is at least as legitimate as rational study. The reflexive dilemma that seems to undermine this claim only arises when we try to justify a priori our empirical (or rational) methods. The argument loses its force if we adopt an evolutionary view of knowledge (e.g., Campbell, 1977, cited in Brewer & Collins, 1981) that recognizes that ideas of merit about science cannot be verified a priori and are ultimately selected by experience from among ideas of varying quality. According to this view, epistemological assumptions are tentative hypotheses rather than a priori truths (cf. Arber, 1954). On the basis of the preceding analysis, we next examine findings and empirically based theory relevant to the observational–theoretical distinction.

A Cognitive Approach to Observational and Theoretical Terms

The analysis that follows is based on a specific cognitive theory and related evidence. The theoretical orientation is that of dual coding theory (Paivio, 1971, 1986), which includes cognitive constructs and mechanisms that are relevant to several of the contentious claims about observational and theoretical terms. A central assumption in the theory is that experiences are represented in concrete, modality-specific ways that reflect the original events on which mental representations are based. A verbal cognitive system represents and processes linguistic information, whereas a nonverbal system handles perceptual-motor information concerning environmental objects and events. The two representational systems are functionally independent, but appropriate experiences produce referential connections between the verbal and nonverbal systems, as well as associative connections among representations within each system. Figure 1 shows the structural features of dual coding theory that are most relevant here.

According to the theory, collections of mental representations are activated directly by external stimuli and indirectly by referential and associative connections. In the case of words or statements, cognitive activity starts at particular verbal representations and spreads throughout the system, as in spreading-activation models of cognitive processing (e.g., Collins & Loftus, 1975). This diffusion produces complex and variable patterns of activation that are determined by prior experiences and the present context. A unique assumption of dual coding theory is that the patterns of mental activation can include both verbal and nonverbal components. The assumption of relatively specific word-like and object-like representations has broad empirical support, including research on instructions to use imagery, reports of spontaneous use of such representations, individual differences in verbal and imaginal abilities or habits, and pertinent attributes of materials (see Paivio, 1971, 1986). Much of the research has examined item attributes that are particularly relevant to the observational–theoretical distinction, namely word concreteness and imagery value.

Dual Coding Theory and the Observability of Scientific Terms

The distinction between observational and theoretical terms follows from the same dual coding assumptions that explain functional differences between concrete and abstract words in natural language. In essence, observational terms (e.g., *bar press* and *eye movement*) are concrete words or phrases that label perceptual objects, properties, or activities. Therefore, they have a high probability of being associated through experience with imaginal representations that render terms referentially meaningful. Theoretical terms (e.g., *image* and *ego*), on the other hand, are abstract words for which verbal associative meaning must carry the major semantic burden because referential meaning is impoverished.

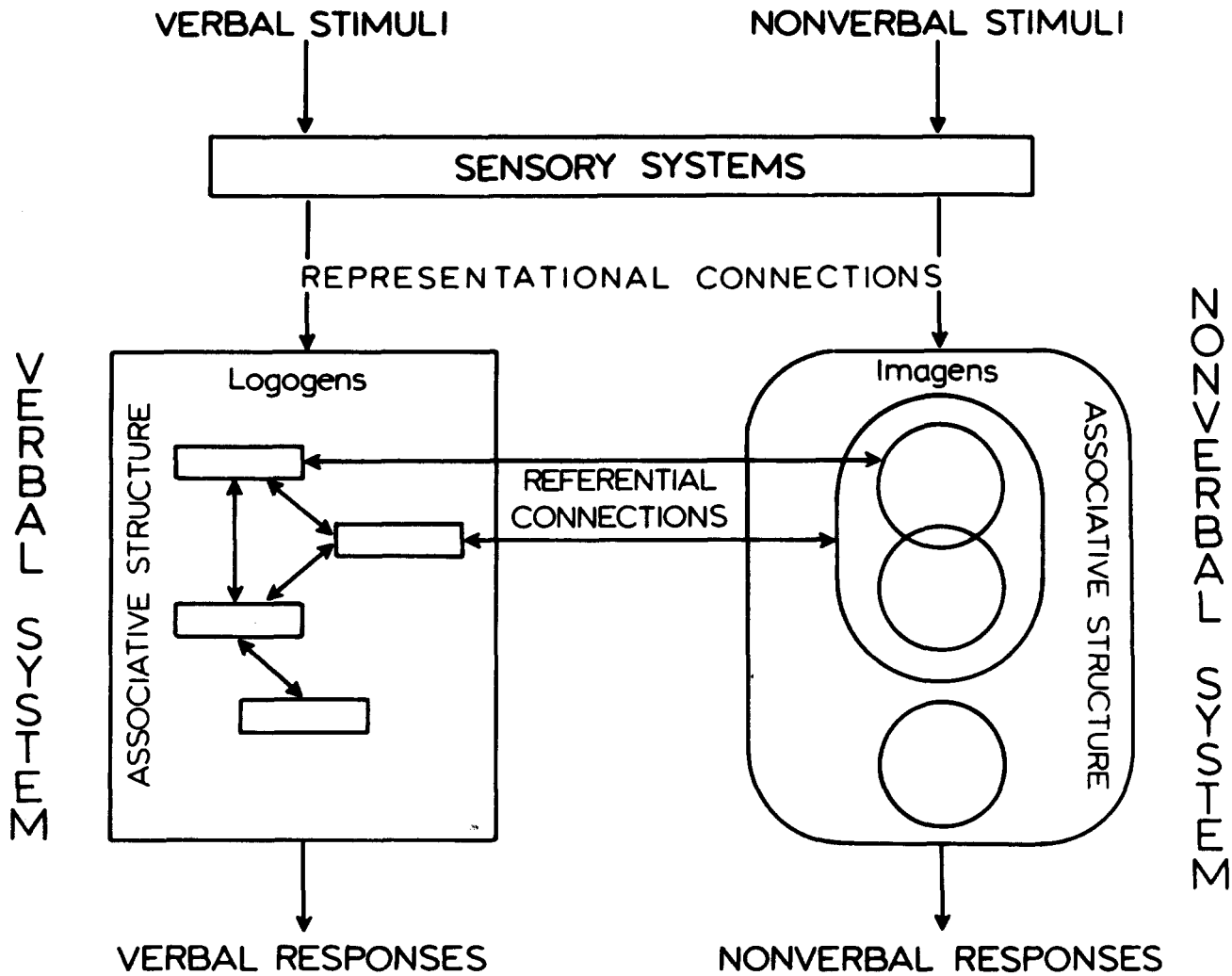
Research on concreteness. If concreteness and observability are similar, then arguments against the observational–theoretical distinction should apply to the distinction between concrete and abstract words, so that the latter distinction would also be difficult to maintain and conceptually problematic. This is not the case. Research has demonstrated the reliability and validity of judgments about word concreteness and imagery value. Paivio, Yuille, and Madigan (1968), for example, had people rate the extent to which words referred to a tangible object, person, or material (concreteness) and the ease or difficulty with which words aroused a mental picture, sound, or other image (imagery value). People readily understand such instructions and produce ratings that are highly consistent both within and across raters (Paivio et al., 1968, p. 6; Toglia & Battig, 1978).

Other research supports the hypothesis that such ratings measure the availability of perceptual knowledge. Rated concreteness and imagery value correlate highly with one another and with such convergent measures as rated tangibility, reaction times for generating images to words (Paivio, 1968), the number of sensory properties elicited by a word (Katz, 1976), and spontaneous reports of imagery experiences. Factor analyses have demonstrated that imagery is independent of theoretically distinct constructs, such as familiarity and meaningfulness (Paivio, 1968). In addition, imagery and concreteness ratings correlate as predicted with performance on memory, comprehension, and other cognitive tasks, even with competing word attributes controlled (Denis, 1979; Paivio, 1968, 1971, 1983a, 1986; Richardson, 1980; Yuille, 1983). This research supports the construct validity of concreteness and imagery ratings and demonstrates that the observability of referents is a meaningful and stable property of words about which there is high intersubjective agreement.

Observability of psychological terms. In studies described more fully elsewhere,[1] we have used cognitive methods to study the observational–theoretical issue in psychology. On the basis of the philosophical literature and feedback from philosophers of science, we defined

[1] An experimental paper describing these studies is in preparation and will be available from the authors in the near future.

Figure 1
Verbal and Nonverbal Symbolic Systems of Dual Coding Theory



Note. From *Mental Representations: A Dual Coding Approach* by A. Paivio, 1986, New York: Oxford University Press. Reprinted by permission. This figure shows the representational units and their referential (between system) and associative (within system) interconnections as well as connections to input and output systems.

observational terms as terms whose meanings were based primarily on direct or instrument-mediated activation of the senses and designated observable entities, and theoretical terms as terms whose meanings were dependent on other terms and referred to hypothesized entities. These definitions were presented to psychology faculty members (19 respondents) who independently nominated 10 observational and 10 theoretical terms from psychology. The 72 terms in Table 1 were selected from those nominated or from glossaries in introductory psychology texts.

Our first study investigated the reliability and validity of measures relevant to the traditional defining attributes of observational and theoretical terms. Groups of psychologists who had not participated in the item generation task rated the terms on several dimensions using seven-point scales. Separate groups rated how observational (n = 9), theoretical (n = 10), or inferential (n = 5) the terms were. The instructions for the observational and theoretical ratings were similar to those used in the generation task. For the inferential scale, subjects assigned high ratings to terms with inferential and derived meanings that could not be known directly. Other psychologists rated the concreteness (n = 7) of the 72 terms using instructions similar to those used in natural language research (e.g., Paivio et al., 1968). That is, subjects were to give high ratings to such tangible entities as objects, materials, or persons and low ratings to abstract concepts that cannot be experienced by the senses.

The results demonstrated that the 72 terms varied consistently in the degree to which they were judged to be observational, theoretical, inferential, and concrete, as

Table 1
Psychological Terms and Mean Observational Ratings

Term	Rating	Term	Rating
Ego	1.44	Aggression	4.44
Image	2.11	Conformity	4.44
Instinct	2.11	Compliance	4.67
Delusion	2.22	Semantic differential	4.78
Egocentrism	2.33	Discrimination	4.89
Schema(ta)	2.33	Peer Group	4.89
Assimilation	2.33	Receptive field	5.11
Consciousness	2.33	Crowding	5.22
Self-concept	2.44	Orienting response	5.22
Meaning	2.44	Labeling	5.25
Representation	2.56	Life events	5.33
Dissonance	2.89	Biofeedback	5.67
Attribution	2.89	Reward	5.67
Perception	2.89	Magnitude estimation	5.78
Encoding	2.89	Threshold	5.78
Trait	3.00	Smile	5.89
Empathy	3.00	Extinction	5.89
Stereotype	3.11	Recall	6.00
Attitude	3.33	Unconditioned stimulus	6.00
Achievement motivation	3.44	Facial expression	6.11
Desensitization	3.56	Touch	6.11
Identity	3.56	Wechsler Adult Intelligence Scale	6.11
Rehearsal	3.67	Proximity	6.25
Stress	3.67	Crying	6.33
Sex roles	3.78	Dopamine	6.33
Anxiety	3.78	Response	6.44
Intelligence	3.78	Serial position	6.44
Modeling	3.88	Speech	6.67
Emotion	3.89	Reaction time	6.78
Attention	3.89	Test score	6.78
Motivation	3.89	Bar press	6.89
Antisocial behavior	4.00	Eye movement	6.89
Learning	4.00	Heart rate	7.00
Adaptation	4.00		
Comprehension	4.11		
Memory	4.11		
Altruism	4.22		
Conflict	4.22		
Psychosis	4.33		

shown here by alpha coefficients of reliability (Table 2), the wide range of the mean observational ratings (Table 1), and the standard deviations of the four scales (Table 2). Furthermore, correlations among the scales (Table 3) confirmed the hypothesis that the four constructs are closely related. The strong, negative correlation between the observational and theoretical measures suggests that they tap opposite poles of a common underlying dimension, and the inferential scale seems to reflect this same construct. Moreover, the three scales derived from the philosophical literature on science correlated highly with concreteness ($rs > .84$), as expected from the dual coding analysis. The common basis of these various measures was shown clearly by a factor analysis on which all four scales loaded .92 or higher on a single factor that accounted for 91.2% of the variability in mean ratings. These results are consistent with natural language research and with the hypothesis that observability provides a meaningful way to classify psychological terms.

The psychologists in our study, most of whom are practicing scientists, made judgments similar to those made by individual scholars in traditional works on the observational–theoretical distinction, so it is interesting to note that our results contrast with some conclusions made by those scholars, presumably on the basis of reason and intuition. For example, the high degree of agreement about the observability of psychological terms contrasts with arguments that the defining features of observational and theoretical terms do not permit unambiguous classification (e.g., Achinstein, 1968). The high negative correlation between observability and theoretical ratings similarly contradicts suggestions that the observational–nonobservational distinction is entirely separate from the nontheoretical–theoretical distinction (e.g., Churchland, 1985).

Reaction time measures and observability. In a second study, we sought convergent evidence that the meanings of observational and theoretical terms differ on the availability of perceptual knowledge about their referents. Natural language research has shown that reaction times on imagery tasks are slower for abstract words than for concrete words despite small differences in verbal associative latencies (Paivio, 1966). Such findings confirm the dual coding hypothesis that concrete and abstract words have differential access to an imaginal system that can be dissociated from verbal components of meaning. To determine whether similar processes operate for psychological terms, reaction times were measured for graduate students in psychology to think of either a related word ($n = 13$) or a mental image ($n = 13$) for each of the 72 terms. We also measured the incidence of spontaneous imagery by later asking subjects in the word association condition to indicate for which words they had experienced a mental image during the association task.

Figure 2 shows mean reaction times as a function of grouped levels of rated observability. Observational and theoretical terms differed primarily on imagery reaction times, which averaged over six seconds at the lowest levels of observability and decreased consistently with in-

Table 2
Descriptive Statistics, Alpha Coefficients, and Factor Loadings for the Rating Scales

Rating Scale	*M*	*SD*	Alpha	Loading
Observational	4.42	1.50	.94	.91
Theoretical	3.84	1.43	.90	−.95
Inferential	3.55	1.44	.82	−.93
Concreteness	4.40	1.25	.90	.94
Consistency	4.24	1.19	.90	.91
Distinctiveness	4.39	1.18	.92	.89
Difficulty	3.22	1.13	.87	−.93
Familiarity	5.18	.78	.62	.33

Table 3
Intercorrelations of Ratings

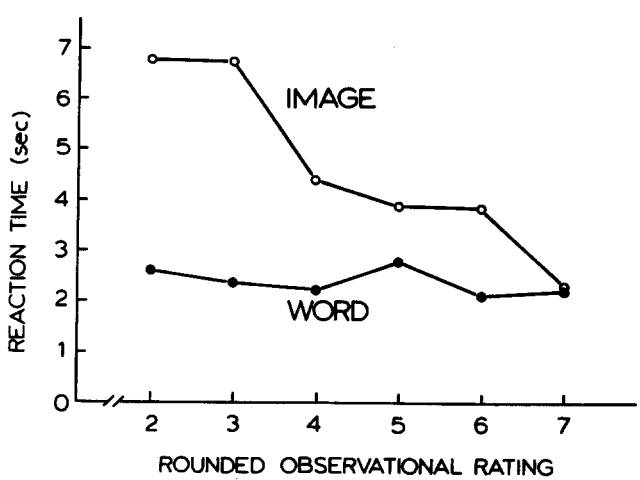
Rating Scale	Observational	Theoretical	Inferential	Concreteness	Consistency	Distinctiveness	Difficulty	Familiarity
Observational		−.89	−.86	.89	.81	.78	−.85	.27
Theoretical			.90	−.90	−.83	−.80	.90	−.32
Inferential				−.85	−.85	−.79	.89	−.28
Concreteness					.84	.84	−.85	.32
Consistency						.92	−.83	.27
Distinctiveness							−.79	.32
Difficulty								−.37
Familiarity								

creases in observability ($r = -.52$, $p < .01$). Reaction times for subjects in the word association task remained relatively constant at about two seconds, and the correlation between verbal reaction time and observability was only −.10. As further evidence for the hypothesized relation between observability and ease of imagery, the number of reported failures to generate images in the imagery condition decreased with increases in the rated observability of the terms ($r = -.31$, $p < .01$), and reports of spontaneous imagery during the word association task increased with observability ($r = .65$, $p < .01$), as shown in Figure 3. Even though their primary task was a verbal one, about one third of the word association subjects reported spontaneous images for observable terms.

These results demonstrate differential access to imaginal knowledge by observational and theoretical terms, and they replicate results of previous studies with concrete and abstract words (Janssen, 1976; Paivio, 1966, 1975). For the words highest in observability, imagery and verbal associative reaction times were equally fast, and the incidence of spontaneous imagery by verbal association subjects was high, despite fast association times and instructions that specifically emphasized verbal associates as opposed to the more general meanings of the terms. That image generation can be so rapid and spontaneous supports the claim that imagery contributes to the meaning of observational terms.

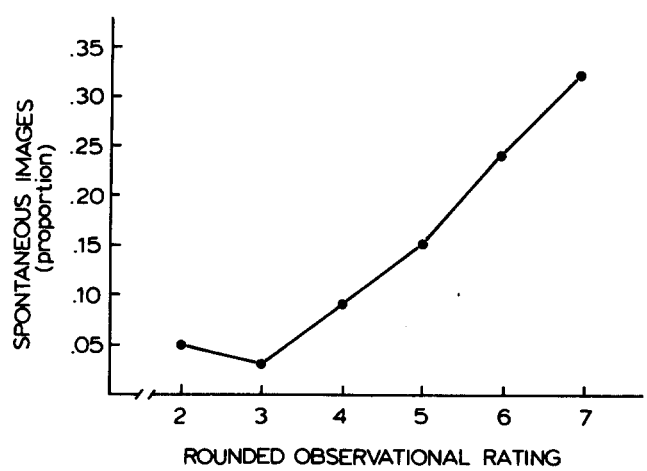
The rating and reaction time results are consistent with the assumption that scientific terms vary systematically and reliably in the directness of their connections to perceptual objects and knowledge, as hypothesized by scholars who support the observational–theoretical distinction. The findings are also consistent with a dual coding view of natural and scientific language. That external referents exist for observational terms also provides some support for the further claim that observational statements are directly verifiable (e.g., Hull, 1943; Nagel, 1961, pp. 83–85). That is, observations of referents and their

Figure 2
Imagery and Word Association Reaction Times for Psychological Terms Varying in Rated Observational Value



Note. There were from 6 to 21 terms per level.

Figure 3
Spontaneous Imagery Reported by Word Association Subjects for Psychological Terms Varying in Rated Observational Value



Note. There were from 6 to 21 terms per level.

relations can confirm or disconfirm statements about hypothetical relations between the observational terms that correspond to the referents. Verification of statements that contain theoretical terms, a primary objective of science, is more inferential because such terms correspond less directly to external referents. We next examined additional claims about observational and theoretical terms.

Relation of Observability to Other Relevant Attributes

Proponents of the observational–theoretical distinction argued that observational terms warrant a special status, not only because they are directly connected to observable referents but also because they are more likely than theoretical terms to show high intersubjective agreement, definiteness of meaning, and other scientifically valued attributes. Scholars who rejected the distinction denied that observational terms have such special properties. The denial followed partly from the assumption that both observational and theoretical terms are equally inferential, but even acceptance of the idea that scientific terms differ in observability does not preclude claims that observational knowledge is less reliable than nonobservational knowledge (e.g., Feyerabend, 1969) or that observational terms are sufficiently inferential to have lost the special properties that proponents of the distinction claimed for them. Because imperfect relations between observability and the disputed attributes of scientific terms are best determined by systematic observation, we examined cognitive findings and theory relevant to the claims.

Many different semantic constructs and even more labels have been used by scholars of science to describe properties of observational and theoretical terms. We have tentatively classified a substantial proportion of the words as referring to either consistency or distinctiveness of meaning, which we refer to collectively as semantic coherence. *Consistency of meaning,* either across people or over time, is one property of scientific terms that has been proposed as a correlate of the observational–theoretical dimension. Such terms as *intersubjective agreement, interpersonal consensus,* and *universal* denote consistency across people, whereas *stability, repeatability, dependability, reliability, replicability,* and similar terms suggest temporal consistency (Meehl, 1978). Advocates of the distinction asserted that the meanings of observational terms are universal because they derive from the objects to which such terms refer (Campbell, 1920/1957; Quine & Ullian, 1978). It was claimed that the meanings of observational terms are consistent over time because phenomena were thought to remain constant despite theoretical changes (Toulmin, 1953). Critics of the distinction argued instead that the meanings of all scientific terms are equally variable and incommensurable (Kuhn, 1962) because they are determined by and change with the theories in which they are embedded (Morick, 1980; Shapere, 1966). Being as heavily influenced by the prevailing worldview as are theoretical terms (Suppe, 1977), the meanings of observational terms are not timeless (Hacking, 1981), highly reliable (Feyerabend, 1969), or permanent (Mahoney, 1976).

Other properties mentioned in the literature were more difficult to classify, but many imply definiteness or *distinctiveness of meaning.* Proponents of the observational–theoretical distinction have asserted that because observational terms can be defined ostensively in terms of referent objects and measurements, they have determinate and definite meanings. The meanings of abstract, theoretical terms, on the other hand, are less definite because they are defined implicitly by theoretical postulates and ultimate uses (Nagel, 1961), may lack complete interpretations (Carnap, 1966), and are learned by elaborate processes related to definition, use, and context (Toulmin, 1953). Scholars who question the observational–theoretical distinction, on the other hand, would emphasize the conditional or context-dependent nature of both observational and theoretical terms, and the general ambiguity and vagueness of all scientific terms.

Consistency and distinctiveness are conceptually related attributes. Being context-sensitive implies not only that a term lacks a single, distinct meaning (i.e., distinctiveness) but also that its meaning will vary as the context changes (i.e., consistency). It also seems probable that variable meanings will often be vague and nondistinctive. For these reasons, consistency and distinctiveness are viewed as components of a more general construct called *semantic coherence* here, and called communicative validity by Begg, Upfold, and Wilton (1978). Not all claims about the properties of observational terms fall neatly into these categories, but many omitted properties (e.g., ease of use, permitting quick decisions) may be secondary effects of semantic coherence. For example, being consistent and definite may make observational terms easy to use. We examined evidence relevant to the semantic coherence of concrete, observational terms in both nonscientific and scientific languages.

Natural language research. Cognitive research on concreteness and semantic coherence provides indirect tests of the hypothesized relation between observability and the consistency and distinctiveness of scientific terms. In a number of studies, concreteness has correlated with a variety of relevant measures. Research on consistency of meaning across persons has shown that different subjects give the same category responses more often for high imagery words than for low imagery words (Kintsch, 1974) and that agreement across subjects in their free associations is also higher for concrete than abstract words (Paivio, 1968). Although stability of meaning has not been studied over extended time periods, the percentage of identical free associations to repeated words is significantly greater for concrete words than for abstract words (Clark, 1978). Reactions to concrete meanings also tend to be more reliable than to abstract meanings when concepts are represented by different words at different times. That is, associations to translations (Kolers, 1963) or to synonyms (Clark, 1978) are more likely to be the same for concrete words than for abstract words. Also related to consistency of meaning, unambiguous concrete words may be less influenced than abstract words by changed contexts (Schwanenflugel & Shoben, 1983).

With respect to distinctiveness or definiteness of meaning, concreteness correlates positively with such relevant measures as rated preciseness (Paivio, 1968) and negatively with measures of confusability, which is the inverse of distinctiveness. Randomly selected abstract words are more likely to share associations, an indicator of confusability, than are concrete words (Paivio & Begg, 1971), and randomly paired abstract words receive higher relatedness ratings than do randomly paired concrete words (Paivio, Clark, & Khan, 1988). Also reflecting the greater confusability of abstract meanings, judgments of sentence comprehensibility are less adversely affected by the random exchange of words between abstract sentences than concrete ones (O'Neill & Paivio, 1978), and abstract homonyms are rated as more similar in meaning than are concrete homonyms (Begg & Clark, 1975).

Other findings that support the hypothesized relation between concreteness and semantic coherence implicate both consistency and distinctiveness of meaning. O'Neill (1972) showed that dictionary definitions were more effective retrieval cues for concrete words than for abstract words. In a related communication task, subjects correctly guessed more concrete than abstract words when given one-word clues generated by other subjects (Begg et al., 1978). Both intersubjective consistency and interitem distinctiveness could contribute to these retrieval differences between concrete and abstract words.

Concreteness is also related to more complex aspects of meaning relevant to philosophical claims about such properties as ease of use. Relative to abstract words, concrete words (a) are rated as easier to define (O'Neill, 1972) and produce better and longer definitions (Reynolds & Paivio, 1968), (b) are rated as easier to categorize (Toglia & Battig, 1978) or make factual statements about (Jones, 1985), and (c) produce a greater number of continuous free associations (Cramer, 1968, p. 75; Lambert, 1955). In some studies, word associations are also faster to concrete words than to abstract words (e.g., Cattell, 1889; Paivio, 1968), but this difference is not always significant. These findings confirm the hypothesis that meanings are more consistent and distinctive for concrete than abstract words. Are similar conclusions justified for scientific terms?

Psychological terms and semantic coherence. New and separate groups of psychology faculty members rated the semantic consistency (n = 10), distinctiveness (n = 12), and difficulty of use (n = 10) of our psychological terms. Subjects were instructed to assign high consistency ratings to terms with stable and reliable meanings that would not vary from one person or time to another. Rated distinctiveness was the degree to which terms had clear and distinct meanings that would not be confused with the meanings of other terms. The difficulty of use scale was based on how difficult it would be to learn, apply, or define the terms. Difficulty was expected to correlate negatively with observability, analogous to the effects just reported for concreteness. We also measured familiarity (n = 9) to determine whether the theoretically motivated scales reflected only general halo effects, and whether differential familiarity could account for the results. The familiarity scale was based on the frequency of experience with the words in print or speech.

The results are shown in Table 2. Individual ratings were very reliable, although somewhat less so for familiarity than for the other scales. Mean ratings on consistency and distinctiveness correlated highly, and both scales correlated negatively with difficulty. The most parsimonious explanation of these correlations is that consistency, distinctiveness, and difficulty reflect a single general construct, such as semantic coherence, or share a strong, common determinant.

The major purpose of this study was to correlate indicators of coherence with the four ratings obtained earlier as measures of observability. Consistent with proponents of the observational–theoretical distinction and research on natural language, the observability and coherence measures were highly correlated ($rs > .77$, see Table 3). A factor analysis of the eight scales produced one factor that accounted for 78% of the variability in the measures. The seven critical scales (excluding familiarity) all had loadings of .89 or higher on this factor (see Table 2). Familiarity had a substantially lower loading, indicating that a general halo effect or differential familiarity does not explain the relations between the variables. Overall, independent groups of psychologists rated the terms essentially the same whether instructed to rate how observational, theoretical, inferential, concrete, consistent, distinctive, or difficult to use the terms are.

Evaluation of the Evidence

Research on both natural and psychological language has produced evidence consistent with the observational–theoretical distinction and with the hypothesized characteristics of observational terms, as proposed by some philosophers of science and as derived from dual coding theory. We briefly consider the generality of these results and some alternative explanations.

Generality

Direct evidence for the observational–theoretical distinction has been presented only for psychology, and the results need to be replicated for other sciences. Nevertheless, we believe that observability is generally relevant to science if only because natural language terms varying in concreteness would play a role in all sciences and because shared cognitive processes would produce similarities between natural and scientific languages. Our conclusions about observational and theoretical terms are also consistent with the following observations from sciences other than psychology.

One source of evidence, although underexploited at present, is scientific communication. Scientists appear to use a high proportion of observational terms (e.g., *recall, test score*) or words that imply observations (e.g., *findings, fact*), at least relative to such disciplines as philosophy. Donald (1986), for example, observed that natural science university courses contained a higher percentage of concrete (i.e., observable) concepts than nonscience courses.

We expect that empirical comparisons of text abstractness across disciplines would also support the hypothesis that scientific writing is relatively concrete. Consistent with this prediction, Cozzens (1985) reported that citations in neuropharmacology emphasized facts more than theory, whereas theoretical constructs were primary in the sociology of science.

The distinction between observational and theoretical terms is closely related to the distinction between observation and theory; thus, evidence that scientists differentiate fact and theory and give priority to empirical findings is relevant to our conclusions. Some scientists explicitly distinguish between hypotheses and sense impressions (e.g., Einstein, 1941/1982, pp. 335–337, 1949/1979, p. 21; Newton, 1672/1953, p. 5) and express greater certainty about sense evidence (e.g., Galileo, 1610/1957, p. 28) and facts (e.g., Darwin, 1877/1969, p. 49) than about speculation and theory. Systematic interviews and ratings similarly show that research chemists justify theories on observational grounds (Gilbert & Mulkay, 1984) and that experimental psychologists value objective evidence more than nonexperimental psychologists (Kimble, 1984). Observation and the relation between observation and theory are stressed in tests of scientific thinking and interests (e.g., Frederiksen & Ward, 1978; Jackson, 1977) and in the stated objectives of science education (e.g., Reif & St. John, 1979). An emphasis on facts may also explain why concrete laboratory experiences are so important in the personal development of scientists (e.g., Einstein, 1949/1979, p. 15; Roe, 1952).

Our dual coding interpretation of the observational–theoretical distinction was based on the assumption of differential access to nonverbal knowledge. This explanation, as well as the generality of the distinction between observational and theoretical terms, is therefore supported by evidence that imagery plays important roles in scientific thinking (e.g., Gruber, 1974; Miller, 1984; Paivio, 1983b; Roe, 1951; Shepard, 1978). Imagery is implicated in mathematics, scientific problem solving, the creative use of models and analogies, and other aspects of science. As further evidence of a reliance on nonverbal processes related to imagery, drawings are used by scientists to explain their ideas in interviews (Gilbert & Mulkay, 1984) and in scientific texts and articles, and nonverbal methods appear to be used more frequently in teaching science than nonscience classes (Donald, 1980). The relevance of concrete experience to the meaning of scientific terms is also suggested by the role of metaphor in science (e.g., Hoffman, 1985).

These diverse findings are consistent with our results for psychological terms and suggest that the distinction between observational and theoretical terms is generally important in science.

Alternative Interpretations

Our findings support the hypothesis that theoretical terms in psychology are less directly connected to observation than are observational terms and have less consistent and definite meanings. Because the gap between observation and theory permits multiple interpretations, however, the hypothesis itself implies that drawing such theoretical inferences is prone to error. We consider briefly some explanations for our results that avoid a theoretical distinction between observational and theoretical terms.

One way to avoid our conclusions is to attribute the results to some other causal variable confounded with observability and its theoretically related constructs. We demonstrated that familiarity could not account for the results, but many potential confounds exist. For example, Achinstein (1968) speculated that judgments about the observational status of scientific terms may actually be based on theoretical terms being scientific or technical (e.g., *field, neutron*) and observational terms not being scientific (e.g., *red, millimeter*). This hypothesis could be tested empirically, but it does not fit well with available data. Many of our most observable psychological terms (e.g., *dopamine, unconditioned stimulus,* and *magnitude estimation*) seem just as scientific as our theoretical terms (e.g., *image, consciousness*) and perhaps more so. Nor does Achinstein's suggestion explain concreteness ratings of natural language and such other findings as the relation between observability and imagery latencies. Nonetheless, potential confounding variables do need empirical investigation by studies analogous to those that have demonstrated concreteness effects in natural language to be independent of correlated characteristics (e.g., Paivio, 1968).

Some of our findings might be explained by hypothesizing that the ratings actually reflect only beliefs of respondents about scientific terms and not actual properties of the terms themselves. That is, even though psychologists might believe that observational terms have more consistent meanings than theoretical terms and give ratings consistent with those beliefs, tasks less sensitive to beliefs (e.g., definition or communication tasks) could show both kinds of terms to be equally variable. Such hypotheses are strained by several aspects of our results. For one thing, psychologists showed remarkable consistency in their ratings of observability and concreteness for individual words. To explain this finding, subjects would require highly specific beliefs about particular words and not just a general conviction in the observational–theoretical distinction. Moreover, results involving reaction time, reports of spontaneous imagery, and other measures are less easily explained by beliefs than are ratings. A further difficulty for the hypothesis of general beliefs is that our findings are consistent with other phenomena of science and with natural language research using objective measures of consistency, distinctiveness, and related traits.

Finally, the results might be reconciled with cognitive theories other than the dual coding view of scientific language that we have presented. Common coding theories of cognition portray mental representations as conceptual entities (e.g., propositions) that are more abstract than their names or other kinds of specific perceptual information (e.g., Potter, Kroll, Yachzel, Carpenter, & Sherman, 1986). Such theories contain no intrinsic mecha-

nism on which to base a distinction between observational and theoretical terms and are more compatible generally with the view that all scientific terms are equally abstract. Nonetheless, common coding theories can be made more similar to dual coding theory by adding assumptions to account for the distinction between scientific terms and the analogous dimension of concreteness. Dual coding theory requires no such modification, having independently predicted many of the findings that support the observational–theoretical distinction. Despite the success of dual coding theory, however, further research is needed to determine which class of theory best explains the collective properties of scientific and natural languages, including the relations reported in this article.

Conclusions

The hypothesis that observational and theoretical terms have distinct qualities relevant to the understanding and practice of science has survived several tests. The observational–theoretical distinction is supported by empirical tests of its validity and of its relation to the semantic consistency and distinctiveness of psychological terms. The distinction is also consistent with relevant evidence from natural language and with a plausible theory of human cognition. Moreover, rational arguments against the distinction are inconclusive, and other phenomena of science suggest that the distinction is generally valid. We tentatively conclude that observational terms refer more directly to observable phenomena than do theoretical terms and are relatively more stable and definite in their meanings.

We also conclude that dual coding theory provides a natural framework for many of the theoretical issues and empirical results relevant to the observational–theoretical distinction. This fit derives from the principled basis in dual coding theory for distinguishing concrete and abstract concepts that have differential access to nonverbal knowledge about the world. Other findings, about which definite predictions were not made, are readily derived from dual coding theory. For example, the evidence for a relation between concreteness and semantic coherence suggests that the unique cognitive structures of concrete and abstract words affect consistency and distinctiveness of meaning. Perhaps semantic coherence is related to the availability of two codes for concrete words but not abstract words, or the imaginal knowledge associated with concrete words may demonstrate more semantic coherence than the predominately verbal knowledge of abstract words.

More generally, we have demonstrated that studies of science and cognitive psychology produce mutual benefits and that the literature on science is a fertile source of theoretical ideas for empirical investigation by cognitive psychologists. Our research examined basic aspects of the observational–theoretical distinction, but scholars of science have also been concerned with how theoretical terms provide explanations for varied phenomena (cf. Nagel, 1961), the order in which observational and theoretical terms evolve (e.g., Cohen, 1980), and how observational and theoretical terms are related by nomological networks (e.g., Feigl, 1970, p. 6; Nagel, 1961; Kerlinger, 1973, p. 34). Such questions about science suggest additional cognitive studies and give a new perspective to related theories of natural language. For example, ideas about nomological networks in science are relevant to psychological theories about category structures and to related natural language findings, such as the apparent correlation between category structures and concreteness (e.g., Kintsch, 1974; Toglia & Battig, 1978).

With respect to the practice of science and metatheoretical criticisms of empirical psychology, the data suggest that scientists do and ought to maintain distinct attitudes toward observational and theoretical terms when thinking about or communicating scientific ideas. Observational terms have more stable and universal meanings, and participate in statements that can be empirically validated by virtue of their concrete referents. More intellectual caution is appropriate when theoretical terms and statements are under consideration. Given the evidence on which these conclusions are based, claims that empirical views of scientific psychology constitute a cult (Toulmin & Leary, 1985) or that a scientific psychology is impossible (Robinson, 1985) are too extreme (see also Broadbent, 1973; Kendler, 1981). Healthy criticism provides an antidote for simplistic beliefs in the absolute infallibility of data, but we should be wary of the extreme view that theoretical ideas are just as firm and reliable as observations. An appreciation of the importance of sound empirical foundations is especially needed for the scientific study of human cognition, which all too easily promotes mentalistic explanations without sufficient concern for observational adequacy (cf. Paivio, 1975, p. 287). The cognitive theory and research reported here suggest that empirical observations provide solid foundations for theories of such natural phenomena as scientific language.

REFERENCES

Achinstein, P. (1968). *Concepts of science: A philosophical analysis.* Baltimore: Johns Hopkins University Press.

Anderson, B. F. (1971). *The psychology experiment: An introduction to the scientific method* (2nd ed.). Belmont, CA: Brooks/Cole.

Arber, A. (1954). *The mind and the eye: A study of the biologist's standpoint.* London: Cambridge University Press.

Begg, I., & Clark, J. M. (1975). Contextual imagery in meaning and memory. *Memory and Cognition, 3,* 117–122.

Begg, I., Upfold, D., & Wilton, T. D. (1978). Imagery in verbal communication. *Journal of Mental Imagery, 2,* 165–186.

Borgen, F. H. (1984). Counseling psychology. *Annual Review of Psychology, 35,* 579–604.

Brewer, M. B., & Collins, B. E. (1981). *Scientific inquiry and the social sciences: A volume in honor of Donald T. Campbell.* San Francisco: Jossey-Bass.

Broadbent, D. E. (1973). *In defence of empirical psychology.* London: Methuen.

Bronowski, J. (1978). *The origins of knowledge and imagination.* New Haven: Yale University Press.

Brown, H. I. (1977). *Perception, theory, and commitment: The new philosophy of science.* Chicago: University of Chicago Press.

Campbell, D. T. (1977). *Descriptive epistemology: Psychological, sociological, and evolutionary.* William James Lectures, Harvard University.

Campbell, N. R. (1957). *Foundations of science.* New York: Dover. (Originally published as *Physics: The elements,* 1920)

Carlsmith, J. M., Ellsworth, P. C., & Aronson, E. (1976). *Methods of research in social psychology.* Reading, MA: Addison-Wesley.
Carnap, R. (1966). *An introduction to the philosophy of science.* New York: Basic Books.
Cattell, J. M. (1889). Mental association investigated by experiment. *Mind, 14,* 230–250.
Churchland, P. M. (1985). The ontological status of observables: In praise of the superempirical virtues. In P. M. Churchland & C. A. Hooker (Eds.), *Images of science* (pp. 35–47). Chicago: The University of Chicago Press.
Churchland, P. M., & Hooker, C. A. (Eds.). (1985). *Images of science.* Chicago: The University of Chicago Press.
Clark, J. M. (1978). *Synonymity and concreteness effects on free recall and free association: Implications for a theory of semantic memory.* Unpublished doctoral dissertation, University of Western Ontario.
Cohen, B. P. (1980). *Developing sociological knowledge: Theory and method.* Englewood Cliffs, NJ: Prentice-Hall.
Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82,* 407–428.
Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings.* Chicago: Rand McNally.
Cozzens, S. E. (1985). Comparing the sciences: Citation context analysis of papers from neuropharmacology and the sociology of science. *Social Studies of Science, 15,* 127–153.
Cramer, P. (1968). *Word association.* New York: Academic Press.
Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52,* 281–302.
Darwin, C. (1969). *The autobiography of Charles Darwin.* New York: W. W. Norton. (Original work published 1877)
Denis, M. (1979). *Les images mentales* [Mental images]. Paris: Presses Universitaire de France.
Donald, J. G. (1980). *Structures of knowledge and implications for teaching* (Report No. 6). Montreal: McGill University.
Donald, J. G. (1986). Knowledge and the university curriculum. *Higher Education, 15,* 267–282.
Einstein, A. (1979). *Autobiographical notes* (P. A. Schilpp, Trans.). Chicago: Open Court Publishing. (Original work published 1949)
Einstein, A. (1982). The common language of science. In *Ideas and opinions* (pp. 335–337). New York: Crown. (Original work published 1941)
Feigl, H. (1970). The "orthodox" view of theories: Remarks in defense as well as critique. In M. Radner & S. Winokur (Eds.), *Minnesota studies in the philosophy of science* (Vol. 4, pp. 3–16). Minneapolis: University of Minnesota Press.
Feyerabend, P. K. (1969). Science without experience. *The Journal of Philosophy, 66,* 791–794.
Flugel, J. C. (1964). *A hundred years of psychology.* London: Methuen.
Fodor, J. (1984). Observation reconsidered. *Philosophy of Science, 51,* 23–43.
Frederiksen, N., & Ward, W. C. (1978). Measures for the study of creativity in scientific problem-solving. *Applied Psychological Measurement, 2,* 1–24.
Galileo, G. (1957). *Discoveries and opinions of Galileo* (S. Drake, Trans.). New York: Doubleday. (Original work published 1610)
Gergen, K. J. (1980). Toward intellectual audacity in social psychology. In R. Gilmour & S. Duck (Eds.), *The development of social psychology* (pp. 239–270). London: Academic Press.
Gergen, K. J. (1986). Elaborating the constructionist thesis. *American Psychologist, 41,* 481–482.
Gilbert, G. N., & Mulkay, M. (1984). *Opening Pandora's box: A sociological analysis of scientists' discourse.* London: Cambridge University Press.
Gruber, H. E. (1974). A psychological study of scientific creativity. In H. E. Gruber & P. H. Barrett, *Darwin on man* (pp. 1–257). New York: E. P. Dutton.
Hacking, I. (Ed.). (1981). *Scientific revolutions.* Oxford: Oxford University Press.
Hanson, N. R. (1958). *Patterns of discovery.* Cambridge, England: Cambridge University Press.
Hempel, C. (1952). *Fundamentals of concept formation in empirical science.* Chicago: University of Chicago Press.
Hoffman, R. R. (1985). Some implications of metaphor for philosophy and psychology of science. In W. Paprotte & R. Dirven (Eds.), *The ubiquity of metaphor* (pp. 327–380). Amsterdam: John Benjamins.
Hull, C. L. (1943). *Principles of behavior: An introduction to behavior theory.* New York: Appleton-Century-Crofts.
Jackson, D. N. (1977). *Jackson vocational interest survey manual.* London, Canada: Research Psychologists Press.
Janssen, W. (1976). *On the nature of the mental image.* Soesterberg, The Netherlands: Institute for Perception TNO.
Jones, G. V. (1985). Deep dyslexia, imageability, and ease of predication. *Brain and Language, 24,* 1–19.
Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80,* 237–251.
Katz, A. N. (1976). *Verbal concept identification: Disentangling the dominance effect.* Unpublished doctoral dissertation, University of Western Ontario.
Kendler, H. H. (1981). *Psychology: A science in conflict.* New York: Oxford University Press.
Kerlinger, F. N. (1973). *Foundations of behavioral research* (2nd ed.). New York: Holt, Rinehart & Winston.
Kerlinger, F. N. (1979). *Behavioral research: A conceptual approach.* New York: Holt, Rinehart & Winston.
Kimble, G. A. (1984). Psychology's two cultures. *American Psychologist, 39,* 833–839.
Kintsch, W. (1974). *The representation of meaning in memory.* Hillsdale, NJ: Erlbaum.
Koch, S. (1981). The nature and limits of psychological knowledge: Lessons of a century qua "science." *American Psychologist, 36,* 257–269.
Kolers, P. A. (1963). Interlingual word associations. *Journal of Verbal Learning and Verbal Behavior, 2,* 291–300.
Krasner, L., & Houts, A. C. (1984). A study of the "value" systems of behavioral scientists. *American Psychologist, 39,* 840–850.
Kuhn, T. S. (1962). *The structure of scientific revolutions.* Chicago: University of Chicago Press.
Kuhn, T. S. (1979). Metaphor in science. In A. Ortony (Ed.), *Metaphor and thought* (pp. 409–419). Cambridge, England: Cambridge University Press.
Lambert, W. E. (1955). Associational fluency as a function of stimulus abstractness. *Canadian Journal of Psychology, 9,* 103–106.
Landesman, C. (Ed.). (1970). *The foundations of knowledge.* Englewood Cliffs, NJ: Prentice-Hall.
Losee, J. (1980). *A historical introduction to the philosophy of science* (2nd ed.). Oxford, England: Oxford University Press.
Mahoney, M. J. (1976). *Scientist as subject: The psychological imperative.* Cambridge, MA: Ballinger.
Mandler, G., & Kessen, W. (1959). *The language of psychology.* New York: Wiley.
Martin, M. (1972). *Concepts of science education: A philosophical analysis.* Glenview, IL: Scott, Foresman.
Meehl, P. E. (1978). Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. *Journal of Consulting and Clinical Psychology, 46,* 806–834.
Messer, S. B. (1985). Choice of method is value laden too. *American Psychologist, 40,* 1414–1415.
Meyers, L. S., & Grossen, N. E. (1974). *Behavioral research: Theory, procedure, and design.* San Francisco: Freeman.
Michalos, A. C. (1970). *Improving your reasoning.* Englewood Cliffs, NJ: Prentice-Hall.
Miller, A. I. (1984). *Imagery in scientific thought: Creating 20th century physics.* Boston, MA: Birkhauser.
Morick, H. (Ed.). (1980). *Challenges to empiricism.* Cambridge, MA: Hackett.
Nagel, E. (1961). *The structure of science: Problems in the logic of scientific explanation.* New York: Harcourt, Brace, & World.
Newton, I. (1953). *Newton's philosophy of nature: Selections from his writings* (H. S. Thayer, Ed.). New York: Hafner Press. (Original work published 1672)
Newton-Smith, W. H. (1981). *The rationality of science.* Boston: Routledge & Kegan Paul.
Nisbett, R., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment.* Englewood Cliffs, NJ: Prentice-Hall.
O'Neill, B. J. (1972). Definability as an index of word meaning. *Journal of Psycholinguistic Research, 1,* 287–298.

O'Neill, B. J., & Paivio, A. (1978). Semantic constraints in encoding judgments and free recall of concrete and abstract sentences. *Canadian Journal of Psychology, 32,* 3–18.
Paivio, A. (1966). Latency of verbal associations and imagery to noun stimuli as a function of abstractness and generality. *Canadian Journal of Psychology, 20,* 378–387.
Paivio, A. (1968). A factor-analytic study of word attributes and verbal learning. *Journal of Verbal Learning and Verbal Behavior, 7,* 41–49.
Paivio, A. (1971). *Imagery and verbal processes.* New York: Holt, Rinehart, & Winston. (Reprinted by Erlbaum, Hillsdale, NJ, 1979)
Paivio, A. (1975). Neomentalism. *Canadian Journal of Psychology, 29,* 263–291.
Paivio, A. (1983a). The empirical case for dual coding. In J. C. Yuille (Ed.), *Imagery, memory and cognition: Essays in honor of Allan Paivio* (pp. 307–332). Hillsdale, NJ: Erlbaum.
Paivio, A. (1983b). The mind's eye in arts and science. *Poetics, 12,* 1–18.
Paivio, A. (1986). *Mental representations: A dual coding approach.* New York: Oxford Press.
Paivio, A., & Begg, I. (1971). Imagery and associative overlap in short-term memory. *Journal of Experimental Psychology, 89,* 40–45.
Paivio, A., Clark, J. M., & Khan, M. (1988). Effects of concreteness and semantic relatedness on composite imagery ratings and cued recall. *Memory & Cognition, 16,* 422–430.
Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology Monograph Supplement, 76*(1, Pt. 2).
Peterson, C. R., & Beach, L. R. (1967). Man as an intuitive statistician. *Psychological Review, 68,* 29–46.
Popper, K. R. (1965). *Conjectures and refutations.* New York: Basic Books.
Potter, M. C., Kroll, J. F., Yachzel, B., Carpenter, E., & Sherman, J. (1986). Pictures in sentences: Understanding without words. *Journal of Experimental Psychology: General, 5,* 31–47.
Pylyshyn, Z. (1979). Metaphorical imprecision and the "top-down" research strategy. In A. Ortony (Ed.), *Metaphor and thought* (pp. 420–436). Cambridge, England: Cambridge University Press.
Quine, W. V. (1951). The two dogmas of empiricism. *The Philosophical Review, 60,* 20–43.
Quine, W. V., & Ullian, J. S. (1978). *The web of belief* (2nd ed). New York: Random House.
Reif, F., & St. John, M. (1979). Teaching physicists' thinking skills in the laboratory. *American Journal of Physics, 47,* 950–957.
Reynolds, A., & Paivio, A. (1968). Cognitive and emotional determinants of speech. *Canadian Journal of Psychology, 22,* 164–175.
Richardson, J. T. E. (1980). *Mental imagery and human memory.* London: Macmillan.
Robinson, D. N. (1985). Science, psychology, and explanation: Synonyms or antonyms. In S. Koch & D. E. Leary (Eds.), *A century of psychology as science* (pp. 60–74). New York: McGraw-Hill.
Roe, A. (1951). A study of imagery in research scientists. *Journal of Personality, 19,* 459–470.
Roe, A. (1952). *The making of a scientist.* New York: Dodd Mead.
Rorer, L. G., & Widiger, T. A. (1983). Personality structure and assessment. *Annual Review of Psychology, 34,* 431–463.
Rosenthal, R. (1966). *Experimenter effects in behavioral research.* New York: Appleton-Century-Crofts.
Royce, J. R. (1982). Philosophical issues, division 24, and the future. *American Psychologist, 37,* 258–266.
Schwanenflugel, P., & Shoben, E. J. (1983). Differential context effects in the comprehension of abstract and concrete materials. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9,* 82–102.
Segal, E. M., & Lachman, R. (1972). Complex behavior or higher mental process: Is there a paradigm shift? *American Psychologist, 27,* 46–55.
Selltiz, C., Wrightsman, L. S., & Cook, S. W. (1976). *Research methods in social relations* (3rd ed.). New York: Holt, Rinehart, & Winston.
Shapere, D. (1966). Meaning and scientific change. In R. G. Colodny (Ed.), *Mind and cosmos: Essays in contemporary science and philosophy* (pp. 41–85). Pittsburgh: University of Pittsburgh Press.
Shapere, D. (1985). Observation and the scientific enterprise. In P. Achinstein & O. Hannaway (Eds.), *Observation, experiment, and hypothesis in modern physical science* (pp. 21–45). Cambridge, MA: MIT Press.
Shepard, R. N. (1978). The mental image. *American Psychologist, 33,* 125–137.
Sklar, L. (1985). Modestly radical empiricism. In P. Achinstein & O. Hannaway (Eds.), *Observation, experiment, and hypothesis in modern physical science* (pp. 1–20). Cambridge, MA: MIT Press.
Smith, E. E., & Medin, D. L. (1981). *Categories and concepts.* Cambridge, MA: Harvard University Press.
Stevens, S. S. (1939). Psychology and the science of science. *Psychological Bulletin, 36,* 221–263.
Suppe, F. (Ed.). (1977). *The structure of scientific theories* (2nd ed.). Urbana: University of Illinois Press.
Symposium on operationism. (1945). *Psychological Review, 52,* 241–294.
Thouless, R. H. (1974). *Straight and crooked thinking.* London: Pan Books.
Toglia, M. P., & Battig, W. F. (1978). *Handbook of semantic word norms.* Hillsdale, NJ: Erlbaum.
Toulmin, S. (1953). *The philosophy of science: An introduction.* New York: Harper & Row.
Toulmin, S., & Leary, D. E. (1985). The cult of empiricism in psychology, and beyond. In S. Koch & D. E. Leary (Eds.), *A century of psychology as science* (pp. 594–617). New York: McGraw-Hill.
Tweney, R. D., Doherty, M. E., & Mynatt, C. R. (Eds.). (1981). *On scientific thinking.* New York: Columbia University Press.
Underwood, B. J. (1957). *Psychological research.* New York: Appleton-Century-Crofts.
van Fraassen, B. C. (1980). *The scientific image.* Oxford, England: Oxford University Press.
van Fraassen, B. C. (1985). Empiricism in the philosophy of science. In P. M. Churchland & C. A. Hooker (Eds.), *Images of science* (pp. 245–308). Chicago: University of Chicago Press.
Weimer, W. B. (1979). *Notes on the methodology of scientific research.* Hillsdale, NJ: Erlbaum.
Wolman, B. B. (1971). Does psychology need its own philosophy of science? *American Psychologist, 26,* 877–886.
Yuille, J. C. (Ed.). (1983). *Imagery, memory and cognition: Essays in honor of Allan Paivio.* Hillsdale, NJ: Erlbaum.