## Toward a Richer View of the Scientific Method

The Role of Conceptual Analysis

Armando Machado *U* Francisco J. Silva *U* 

Universidade do Minho University of Redlands

Within the complex set of activities that comprise the scientific method, three clusters of activities can be recognized: experimentation, mathematization, and conceptual analysis. In psychology, the first two of these clusters are well-known and valued, but the third seems less known and valued. The authors show the value of these three clusters of scientific method activities in the works of the quintessential scientist Galileo Galilei. They then illustrate how conceptual analysis can be used in psychology to clarify the grammar and meaning of concepts, expose conceptual problems in models, reveal unacknowledged assumptions and steps in arguments, and evaluate the consistency of theoretical accounts. The article concludes with a discussion of three criticisms of conceptual analysis.

*Keywords:* conceptual analysis, criticism, Galileo, scientific method, theory

odern science is characterized by a rich variety of activities. Researchers observe and experiment in the field and laboratory, devise new techniques and refine old ones, invent and elaborate concepts, quantify functional relations between variables, build models and theories, debate controversial issues, write and review manuscripts and grant proposals, and so on. Within this set of activities we can identify clusters or groups of related activities. Observation/experimentation comprises the actions researchers perform to generate theories and test hypotheses. This cluster includes operationally defining variables, manipulating and controlling them, implementing a research design, selecting research participants and allocating them to groups, and conducting the experiment itself. Quantification/mathematization consists of analyzing data and formulating laws, models, and theories on the basis of empirical findings and previous knowledge. Other activities in this cluster include discovering the mathematical function that may link two variables, finding the bestfitting parameters of the function using statistical techniques, integrating two or more quantitative functions into higher order laws, and testing the statistical significance of the relation between variables. We recognize in these two clusters-experimentation and mathematization, for short-most of the activities subsumed under the broad conception of scientific methods in psychology (e.g., Slife & Williams, 1995; Stanovich, 1998). To a large extent, these two clusters match the activities described in most textbooks and courses on research methods (experimentation) and applied mathematics and statistics for psychologists (mathematization).

Although experimentation and mathematization are well-known components of the scientific method and their importance is unquestionable-in fact, psychologists have acknowledged their importance since the very beginnings of their science in the last quarter of the 19th century (Burnham, 1987)-there is a third cluster of activities that also is an essential component of the scientific method. Theoretical/conceptual analysis comprises the actions researchers engage in when they evaluate the language of their science. These actions include but are not limited to assessing the clarity or obscurity of scientific concepts, evaluating the precision or vagueness of scientific hypotheses, assessing the consistency or inconsistency of a set of statements and laws, and scrutinizing arguments and chains of inferences for unstated but crucial assumptions or steps. In general, then, the goal of conceptual analysis is to increase "the conceptual clarity of a theory through careful clarifications and specifications of meaning" (Laudan, 1977, p. 50).

Although psychologists undoubtedly engage in conceptual analysis, this aspect of the scientific method does not seem as well-known or as acknowledged as experimentation and mathematization (see Machado, Lourenço, & Silva, 2000). First, whereas most introductory psychology textbooks include a chapter on research methods and basic statistical ideas, and most instructors of introductory courses cover this material within the first week or two of classes, there is no comparable chapter or even sections of a chapter devoted to conceptual analysis. Indeed, in contrast to *methods*, *research methods*, *research design*, *statistics*, *statistical analysis*, *data analysis*, and the like, there is no term or set of terms in introductory psychology or research methods textbooks to describe conceptual analysis. The only related topics that are occasionally mentioned

Armando Machado was funded by a grant from the Portuguese Foundation for Science and Technology. We thank Susan Goldstein for sharing her expertise about construct validity.

Correspondence concerning this article should be addressed to Armando Machado, Instituto de Educação e Psicologia, Universidade do Minho, Campus de Gualtar, 4710 Braga, Portugal or Francisco J. Silva, Department of Psychology, University of Redlands, P. O. Box 3080, 1200 East Colton Avenue, Redlands, CA 92373-0999. E-mail: armandom@iep.uminho.pt or francisco\_silva@redlands.edu



#### Armando Machado

are critical thinking and logic. Second, research methods and statistics courses are universally required of psychology majors and graduate students, but few if any psychology courses on conceptual analysis exist or are required. Third, conceptual analysis is rarely the subject matter of conference presentations, articles, chapters, or books for experimental psychologists. Fourth, when psychologists are asked about the most important concepts or topics that undergraduates should learn, they respond "research methods," "independent and dependent variables," and the like (see Boneau, 1990); conceptual analysis or anything like it is unmentioned. Fifth, methodological and statistical flaws-not conceptual problems-are the reasons most often invoked for rejecting papers submitted for publication (e.g., Peters & Ceci, 1982; see also Broad & Wade, 1982; Ghiselin, 1989; Harcum & Rosen, 1993). Although any of these facts can be interpreted in other ways (e.g., methodological and statistical flaws are "fatal flaws" and therefore are more likely than conceptual flaws to be invoked as reasons for rejecting a manuscript; introductory psychology textbooks are meant to provide an overview of the major subject areas of psychology), it is difficult to explain this constellation of facts without concluding that conceptual analysis is, at least to some degree, devalued in psychology (Machado et al., 2000; Slife & Williams, 1995).

In the first part of this article, we show the importance of conceptual analysis in the work of the quintessential scientist of the modern era, Galileo Galilei. We show that the Italian scientist, often portrayed as personifying the epitome of experimentation and mathematization, used conceptual analysis to formulate clear concepts and testable explanations of the physical world and to reject their opposites. In the second part of the article, we discuss conceptual analysis with reference to examples in psychology. Throughout our presentation of the two sets of examples, we adopt a concrete perspective to illustrate various aspects of conceptual analysis and its targets, to identify recurrent themes across distinct domains, and to describe the limits of the activity. In the third and final part, we adopt a more abstract perspective and address a few criticisms of conceptual analysis.

### Galileo's View of the Scientific Method: Controlled Experiments, Mathematical Analysis, and Conceptual Investigations

Galileo Galilei (1564–1642) was a leading figure of the scientific revolution. Some remember his polemics with the Church concerning the relative merits of the Copernican and Ptolemaic systems and his trial by the Inquisition; others remember his discoveries of the trajectory of projectiles and the swinging of the pendulum or his refinements of the telescope. For physicists, perhaps his major accomplishments were the experiments he performed and the theories he developed on free-falling bodies near the surface of the Earth. We begin with a brief review of that for which Galileo is most well-known (i.e., his experiments and mathematical discoveries) and then illustrate in some detail that for which he is less well-known (i.e., his conceptual analyses in physics and astronomy).

### **Experimentation**

Controlled experiments are a legacy of the scientific revolution. In Galileo's descriptions of his experiments, it is easy to recognize some of the main characteristics of the modern scientific method, namely, manipulating the independent variable, controlling extraneous variables, measuring the dependent variable, and addressing measurement reliability. Consider the following description of Galileo's famous studies with inclined planes.

A piece of wooden moulding or scantling, about 12 cubits long, half a cubit wide, and three finger-breadths thick, was taken; on its edge was cut a channel a little more than one finger in breadth; having made this groove very straight, smooth, and polished, and having lined it with parchment also as smooth and polished as possible, we rolled along it a hard, smooth, and very round bronze ball. [Galileo wanted to study the interrelation between distance traversed and time taken by a falling body, an interrelation not contaminated by friction. Hence his attempt to reduce as much as possible the influence of this "extraneous" variable.] Having placed this board in a sloping position, by lifting one end some one or two cubits above the other, we rolled the ball, as I was just saying, along the channel, noting, in a manner presently to be described, the time required to make the descent. [Because the motion of free-falling bodies occurs very quickly-and therefore is hard to characterize-Galileo slowed their motion by means of the inclined plane.] We repeated this experiment more than once in order to measure the time with an accuracy such that the deviation between two observations never exceeded one-tenth of a pulse-beat. Having performed this operation and having assured ourselves of its reliability, we now rolled the ball only one-quarter the length of the channel.... Next we tried other distances, comparing the time for the whole length with that for the half, or



#### Francisco J. Silva

with that for two-thirds, or three-fourths, or indeed for any fraction. [Galileo varied systematically the independent variable distance—and measured repeatedly the dependent variable—time of descent.].... For the measurement of time, we employed a large vessel of water placed in an elevated position; to the bottom of this set was soldered a pipe of small diameter giving a thin jet of water, which we collected in a small glass during the time of each descent, whether for the whole length of the channel or for a part of its length; the water thus collected was weighed, after each descent, on a very accurate balance; the differences and ratios of the weights gave us the differences and ratios of the times, and this with such accuracy that although the operation was repeated many, many times, there was no appreciable discrepancy in the results. (Galilei, 1638/1914, pp. 178–179)

### **Mathematization**

Another legacy of the scientific revolution is the attempt to formulate mathematically the relation between independent and dependent variables. To illustrate, let us return to Galileo's use of inclined planes. After many experiments of the sort described above, some astute thought experiments, a few false starts, and several decades of thinking about free-falling motion, Galileo arrived at the correct mathematical formulation of the law: "In such experiments, repeated a full hundred times, we always found the spaces traversed were to each other as the squares of the times, and this was true for all inclinations of the plane" (Galilei, 1638/1914, p. 179). Galileo had discovered that  $d_1/d_2 =$  $t_1^2/t_2^2$  or, in modern terminology, that distance is proportional to the square of time,  $d = k \times t^2$ . For some historians of science, the coupling of mathematical reasoning with controlled experiments and observations was Galileo's chief contribution to the development of scientific method in the 17th century (Gower, 1997, p. 37).

### **Conceptual Analysis**

But Galileo's scientific method consisted of more than controlled experiments and mathematics. There was also concern for the clarity of concepts, the coherence of hypotheses, and the consistency of theoretical arguments. In several passages of Galileo's writings, one can see him analyzing the language of science, that is, conducting conceptual investigations. These investigations may be what Stillman Drake (1990), arguably Galileo's best scientific biographer, had in mind when he mentioned "reasoning" alongside observation and mathematics as the three fundamental aspects of Galileo's approach:

[Galileo] agreed that in order to become science, philosophy must throw out blind respect for authority; but he also saw that neither observation, nor reasoning, nor the use of mathematics could be thrown out along with this. True philosophy had to be built upon the interplay of all three, and no combination could supply the absence of any one of them. (Drake, 1957, pp. 223–224)

Next, we examine examples of conceptual analysis in Galileo's work. In each case, one can see the Italian scientist exposing a problem or difficulty associated with a concept, hypothesis, or account. The sources of the problems seem to fall into distinguishable categories.

**Inappropriate or illogical classification.** A major aim of conceptual analysis is to identify conceptual problems, which often arise from inappropriate or illogical classification of concepts. In his scientific manifesto "The Assayer," Galileo provided the following straightforward example when addressing Sarsi, his opponent, about the nature of comets and the functioning of the telescope:

If we go back to examine his [Sarsi's] argument more closely, we find it to be defective because it takes as absolute that which must be understood relatively, or as bounded that which is unbounded. In a word, Sarsi has created an incomplete dichotomy (as logicians call this error) when he divided visible objects into "far" and "near" without assigning limits and boundaries between these. He has made the same mistake as a person who should say, "Everything in the world is either large or small." This proposition is neither true nor false, and neither is the proposition "objects are either near or far"... In order to avoid equivocation Sarsi needed to give his classification at least three parts... Nor should he even stop there; he should give an exact determination of this limit, saying for example: "I call 'medium' a distance of one league; 'far,' that which is more than one league; and 'near,' that which is less." (Galilei, 1623/1957, p. 249)

The point of the example is not whether Galileo's theories were right—in fact, his theory of comets was wrong—but that a particular argument was conceptually defective for the reasons he identified. Galileo's remark that the proposition "Everything in the world is either large or small" is neither true nor false but equivocal is a hall-mark of conceptual analysis.

**Excessively vague and ad hoc explanations.** In this next example, Galileo took issue with Sarsi's excessively vague and ad hoc explanation to account for the paths of comets. When Sarsi remarked that Galileo's account on the nature of comets failed to consider various alternative motions, particularly motions along irregular paths, Galileo pointed out that the concept "irregular path" could not be specified, and he concluded as follows:

Lines are called regular when, having a fixed and definite description, they are susceptible of definition and of having their properties demonstrated. Thus the spiral is regular, and its definition originates in two uniform motions, one straight and the other circular... Irregular lines are those which have no determinacy whatever, and are indefinite and casual and hence undefinable.... Hence to say, "Such events take place thanks to an irregular path" is the same as to say, "I do not know why they occur." The introduction of such [irregular] lines is in no way superior to the "sympathy," "antipathy," "occult properties," "influences," and other terms employed by some philosophers as a cloak for the correct reply, which would be: "I do not know." That reply is as much more tolerable than the others as candid honesty is more beautiful than deceitful duplicity. (Galilei, 1623/1957, p. 241)

Galileo's sensitivity to the definition of the key terms in the argument (regular and irregular lines) is another expression of his keen ability for conceptual analysis.

**Irrefutable hypotheses.** In this third example, Galileo rejected an irrefutable hypothesis. Sarsi wanted him to believe that the Babylonians cooked their eggs by whirling them in slings. Galileo claimed he could not obtain the effect experimentally, but apparently this was insufficient to change Sarsi's mind. Galileo remarked,

All too prudently you [Sarsi] have secured your position by saying that "there is needed for this effect violent motion, a great quantity of exhalations, a highly attenuated material, and whatever else conduces to it." This "whatever else conduces to it" is what beats me, and gives you a blessed harbour, a sanctuary completely secure. (Galilei, 1623/1957, p. 273)

**The nominal fallacy (Naming is not explaining).** When we use a familiar name to label an unfamiliar phenomenon, our understanding of that phenomenon seems to increase automatically. Galileo was aware of this cognitive illusion and joked about it in the following passage:

[Sarsi believed that comets are planets or quasi-planets but Galileo found his arguments fallacious and concluded:] I am not so sure that in order to make a comet a quasi-planet, and as such to deck it out in the attributes of other planets, it is sufficient for Sarsi or his teacher to regard it as one and so name it. If their opinions and their voices have the power of calling into existence the things they name, then I beg them to do me the favor of naming a lot of old hardware I have about my house, "gold." (Galilei, 1623/1957, p. 253)

Similarly, when Aristotle's representative in Galileo's dialogues, Simplicio, stated that everyone knows that gravity causes bodies to fall downwards, Galileo's spokesman, Salviati, replied,

You are wrong, Simplicio; what you ought to say is that everyone knows that it is called "gravity." What I am asking you for is not the name of the thing, but its essence, of which essence you know not a bit more than you know about the essence of whatever moves the stars around. I accept the name which has been attached to it and which has been made a familiar household word by the continual experience that we have of it daily. But we do not really understand what principle or what force it is that moves stones downward, any more than we understand who moves them upward after they leave the thrower's hand, or what moves the moon around. (Galilei, 1632/2001, p. 272)

Unjustified extension of a familiar concept to an unfamiliar domain. Our familiarity with a set of concepts, usually from everyday speech, breeds contempt for their scientific analysis. "Why worry with analysis if we all know what these concepts mean?" one might ask. Frequently, though, such "uncritical assumptions of mutual understanding" (Quine, 1936, p. 90) give us only unjustified feelings of accomplishment. The problem is particularly acute when familiar terms are extended to novel domains, as the following interchange reveals. Simplicio knew how to compare two collections of objects and conclude that Collection A is greater than, equal to, or less than Collection B. He also knew that a line segment has an infinite number of points. But when he saw two line segments, one longer than the other, he attempted to determine which line had more points. That is, he attempted to compare infinite sets and, naturally, he extended to this unfamiliar and new domain the familiar process of comparing finite sets:

Since it is clear that we may have one line greater than another, each containing an infinite number of points, we are forced to admit that, within one and the same class, we may have something greater than infinity, because the infinity of points in the long line is greater than the infinity of points in the short line. This assigning to an infinite quantity a value greater than infinity is quite beyond my comprehension. (Galilei, 1638/1914, p. 31)

Salviati identified the source of Simplicio's puzzlement: "The attributes 'equal,' 'greater,' and 'less,' are not applicable to infinite, but only to finite quantities" (Galilei, 1638/1914, pp. 31–33). The lesson is, when a concept is extended to an unfamiliar domain, one must check which properties of the concept hold in the new domain.

Semantic ambiguity (Similar words may hide different conceptions). If all concepts are theory-laden, then it is imperative to check whether two usages of the concept presuppose the same underlying "theory," lest the door be open to misinterpretations and fruitless disputes. As shown in the next example, Galileo knew that the same word might hide different conceptions.

According to Sarsi, Galileo's idea that friction is the cause of heat is consistent with Aristotle's claim that motion is the cause of heat. Sarsi reasoned as follows: Although motion, as motion, is not the cause of heat, friction cannot occur without motion; therefore, motion is at least derivatively the cause of heat. To which Galileo replied,

But if that is what Aristotle meant, then why didn't he say "friction"? When a man can say definitely what he means by using a simple and appropriate word, why employ an inappropriate one that requires qualification and ultimately becomes transformed into something quite different? But assuming that this was Aristotle's meaning, it still differs from Guiducci's [Galileo's student and spokesman]; for to Aristotle any rubbing of bodies would suffice, even of tenuous ones or of the air itself, whereas Guiducci requires two solid bodies, for he considers that trying to pulverize the air is as great a waste of time as grinding water in the proverbial mortar. (Galilei, 1623/1957, p. 266)

Note that Galileo was not arguing about experimental or quantitative issues, nor was he saying that motion does not cause heat. Instead, he was commenting on concepts, on the clarity or obscurity of their use, and on whether the definitions of two concepts overlapped.

Galileo's works provide several other instances of conceptual analysis dealing with, for example, the perils of reasoning by analogy, ad hoc hypotheses, confusing abstract and concrete terms, and the meaninglessness of concepts that express judgments of value (e.g., a circle is the noblest of shapes). From these examples at the dawn of modern science we see experiments, mathematics, and conceptual analysis all playing a significant role in the advancement of knowledge.

# Examples of Conceptual Analysis in Psychology

In the examples that follow, we provide a concrete, diversified, example-based view of conceptual analysis in psychology and continue to identify some of its fundamental characteristics. We begin with an illustration of how conceptual analysis clarified the grammar and meaning of a concept and thereby solved a conundrum—how an inverted retinal image yields a righted percept. We then illustrate how conceptual analysis can be used to expose problems in models, reveal key assumptions and obscured steps in arguments, and evaluate the consistency of a theoretical account.

# Clarifying the Grammar and Meaning of Concepts

One of the main achievements of the scientific revolution bears directly on psychological matters and provides a nice illustration of how conceptual analysis can resolve a longstanding problem. In 1604, Johannes Kepler correctly described how the image of an object is formed in the retina, a description that is the basis of our modern understanding of vision. Kepler's geometrical analysis assumed that countless cones of light extend from every point on an object (the apex) to a common base on the lens of the eye; as each light ray enters the eye, it is refracted by the cornea and lens before exciting the retina. As a consequence of optics, "that which is to the right outside is depicted on the left on the retina, that to the left on the right, that above below, and that below above" (Kepler, cited in Herrnstein & Boring, 1965, p. 94). But if the retinal image is inverted, how do we see objects upright in their "correct" orientation? The question had disturbed previous researchers and forced them to abandon any account that resulted in an inverted image. For example, Leonardo da Vinci compared the eye to a camera obscura but arranged the optics to preserve an erect image as a necessary condition for upright vision (see Crombie, 1964).

The answer to the "problem" was provided by William Molyneux when he clarified the *grammar* of the concept of

seeing-it is the person who sees, not his or her eyes-and the meaning of concepts such as erect and inverted, up and down, closest to and farthest from the center of the Earth. His argument can be recast as follows: Because (a) an eye does not see the retinal image and (b) the retinal image preserves the relations among the objects, it follows that (c) there is no need to force the image to accord with perception. The problem of how objects are seen upright if their retinal images are inverted is not to be solved empirically but to be dissolved conceptually. Two centuries after Molyneux, the physiologist Volkmann put it thus, "The most natural explanation of upright vision is that it does not require an explanation" (cited in Wade, 1998, p. 325). Gregory (1996) not only pointed out that there is no inverted image problem to solve because the image is not an object that is seen but added the idea that if retinal images were seen, then there would be an infinite regress of images, images on eyes, images of images on eyes, and so on.

### Exposing a Limitation or Problem in a Model

The next example shows how conceptual investigations may target even elegant and sophisticated mathematical theories. In this case, the investigation reveals a problem in the dominant model of animal and human interval timing, scalar expectancy theory, or SET (Gibbon, 1977, 1981; see also Gallistel, 1990). To begin, consider the following discrimination task: A subject (animal or human) is rewarded for choosing a red key after a short stimulus (e.g., a 2-s tone) and a green key after a long stimulus (e.g., an 8-s tone). After some trials, the subject reliably chooses the red and green keys when the stimuli are short and long, respectively. To explain how the subject learns the temporal discrimination, SET postulates an internal clock consisting of a pacemaker that generates pulses, an accumulator that counts the pulses emitted during the stimulus, and two memory stores that save the counts in the accumulator at the end of the short and long signals. Although SET has additional features, for present purposes it is sufficient to note that the temporal discrimination depends on the formation of two distinct memory stores.

But how does the timing system decide, so to speak, into which memory to store the count obtained after a particular stimulus? There is only one plausible answer: When the subject chooses the red key and is rewarded for this response, the count in the accumulator is saved to one memory store (call it A); when the subject chooses the green key and is rewarded for this choice, the count in the accumulator is saved into the other memory store (call it B). Thus, counts are saved to a particular memory store on the basis of the structural features of the task (e.g., choosing this or that distinctive key and getting a reward). The theory has no major difficulty accounting for the discrimination in this task.

Consider now a second task. A pigeon receives food for pecking a key after either 10 s or 120 s have elapsed since the onset of the trial. No cue signals whether the current trial will be short or long, and the two types of trials are equally probable. The result of this training is that during the long trials, the pigeon's average rate of pecking increases from the beginning of the trial until approximately 10 s have elapsed, at which point the rate of pecking decreases before beginning to increase again and peaking at 120 s (e.g., Catania & Reynolds, 1968; Leak & Gibbon, 1995). SET explains this pattern of responding by assuming that the animal stores the counts obtained at 10 s and 120 s into distinct memory stores (just as in the first example above). As Leak and Gibbon (1995, p. 6) put it, "In SET, there is assumed to be a single clock but an independent memory distribution for each criterion time interval."

But how does the timing system decide into which memory to store the count obtained after a particular interval? The answer offered for the first example will not work here because there is only one key and no distinct signal to cue the two trials. That is, the counts cannot be "directed" to and saved into different memory stores on the basis of the structural features of the task. SET explains the temporal discrimination by assuming that small counts are saved in Store A and large counts are saved in Store B. But this explanation is a classic case of begging the question (petitio principi for logicians): That which was supposed to be explained by the theory, namely, the discrimination between two intervals or its equivalent (the discrimination between two sets of counts) was assumed in the explanation-the system knows the difference between "small" and "large" counts, which is equivalent to the difference between short and long intervals. Any explanation that assumes the existence of distinct memory bins begs the question of how they were initially formed.

# Revealing Unacknowledged Assumptions and Steps in Arguments

Conceptual analysis may help reveal unacknowledged but fundamental assumptions or steps in an argument or chain of inferences. Consider the following example. A 5-monthold baby is shown a doll. Next, an opaque screen rotates up to hide the doll. Then the baby sees a second doll being put behind the screen. Finally, when the screen is dropped, the baby sees either two dolls (the so-called possible result) or only one doll (the so-called impossible result). The experimenter records the total time the baby looks at the resulting doll(s). The typical finding is that babies stare longer at the impossible outcome than at the possible outcome (e.g., Wynn, 1992, 1995). Similar experiments with cotton-top tamarin monkeys and dogs (substituting food rewards for dolls) have yielded similar results (e.g., Uller, Hauser, & Carey, 2001; West & Young, 2002).

These results have been explained as follows: Babies, dogs, and monkeys (a) compute the outcome of certain operations (e.g., 1 + 1 = 2) and (b) on the basis of that computation, predict what is behind the screen; (c) when the screen drops, they compare the number of items they expected to see with the number that they actually see; (d) when the comparison yields a mismatch (e.g., 2 vs. 1) they are surprised and therefore stare longer at the unexpected outcome. As Wynn (1992) put it,

[When the experimenter placed the second item behind the screen, the] infants could clearly see the nature of the arithmetical oper-

ation being performed, but could not see the result of the operation.... Infants look longer at unexpected events than expected ones, thus, if they are able to compute the numerical results of these arithmetical operations, they should look longer at the incorrect than at the correct results. (p. 749)

But note that by treating the act of bringing a second doll to the stage as an arithmetic operation, the physical action performed by the experimenter and the mental operation executed by the baby are treated as one and the same. Other authors have treated actions and mental operations similarly: "The results suggest that the dogs were anticipating the outcome of the calculations they observed, thus suggesting that dogs may have a rudimentary ability to count" (West & Young, 2002, p. 183, italics added). Although infants and animals might perform computations and calculations upon seeing the experimenter's actions, computations and calculations are not themselves seen. By envisaging the experimenter's actions as arithmetic operations or by stating that dogs observe calculations (as opposed to *perform* calculations), two distinct activities (the experimenter's actions and the subject's purported calculations) were conceptualized as a single one. Perhaps what the authors assumed was that the experimenter's action of adding a doll or a food reward triggered the subject's mental operation of addition. Although this assumption may be correct, it needs to be made explicit and analyzed empirically before we can discover the properties of the mapping between the experimenter's action and a subject's mental operation (e.g., What roles do temporal and spatial attributes of the experimenter's actions, the trajectories of the dolls or food rewards, and the number of times each doll or food reward is presented play in how the child, monkey, or dog comes to know that the physical operation was the occasion for the arithmetic operation?).

### **Questioning the Consistency of an Account**

The next example is potentially more controversial than the preceding ones. We include it to illustrate additional aspects of conceptual analysis—in this case, how it questions the consistency of an account that uses concepts from different universes of discourse: mental intentions, neural representations, and motor behaviors. The use of different languages, as it were, is a perennial source of conceptual difficulties. We also want to show that conceptual analysis may produce more questions than it answers, stimulate spirited debate, and inspire further experiments.

An article published in the prestigious journal *Science* opened with the following paragraph:

A motor action is voluntary if and only if it is intended. William James put forward the ideomotor theory of action, which states that any intention or idea of an action has the tendency to cause the relevant movements. To prevent ourselves from committing action errors, it is frequently important that we attend to our intentions before executing an action. (Lau, Rogers, Haggard, & Passingham, 2004, p. 1208)

The authors then described an experiment designed to study the neural correlates of attention to the intention to perform an act. Using functional magnetic resonance imaging (fMRI), they compared two conditions in a withinsubjects design. In one condition the participants pushed a button while paying attention to the act itself; in the other condition the participants pushed the button while paying attention to the intention to push the button. Thus "the only difference between the two conditions was the focus of attention. This was manipulated by requiring them to report either the time at which they felt the intention to move or the time at which they actually made the movement" (Lau et al., 2004, p. 1208). The results showed that whereas the act was reported about 30 ms before it actually took place, the intention to act was reported about 230 ms before the act. In addition, the fMRI revealed "specific activations associated with attention to intention" (p. 1209), specifically, an enhancement of activity in the pre-supplemental motor area (pre-SMA) coupled with activations in the right dorsal prefrontal cortex. The authors concluded that activity in the pre-SMA reflects the representation of intention.

As interesting and provocative as these results are, the scientific game requires us, as scientists, to analyze the authors' arguments and evaluate their conclusions. Should we accept Lau et al.'s assertions that what people reported at 230 ms was their intention to act, that these intentions are represented in the pre-SMA, and that they are causally related to the behavior of pushing the button? The answers require conceptual analysis, a reasonably detailed example of which follows.

**Identify the purpose and structure of the argument.** The experiment was designed to show that attending to the intention of an act has specific brain correlates and that these correlates differ from the correlates of attending to the act itself. More generally, Lau et al. (2004) were searching for the neural foundations of voluntary or intentional behavior. Their argument has at least three components in the form of implicit or explicit assumptions: (a) Intention is a mental representation with a neural instantiation (e.g., increased activity in a particular brain region), and thus intention is an occurrence, something that can be the target of "effective editing and evaluation" (p. 1210); (b) intention precedes and efficiently causes voluntary behavior; and (c) given that an intention is an occurrence, attention can be focused on it.

Critique the structure of the argument. We mention only critiques to each part separately and ignore potential critiques of their interconnection. First, if an intention is a mental or neural occurrence, then we can ask: "What causes an intention?" If we reply that an intention is caused by "something" intentional, then we have an infinite regress because Intention A would be caused by Intention B, which would be caused by Intention C, and so on. But if we reply that intentional behavior is caused by something unintentional, then in what sense is the intentional behavior voluntary? Second, if intentions efficiently cause voluntary behavior, then the behavior should occur once it is intended. But this seems unlikely. A tennis player's intention to hit a cross-court forehand on the next shot does not mean that he will; a return sent to his backhand side, for example, may cause him to hit a backhand down the line. Third, we invite readers to imagine

how they would respond to three different instructions: "Pay attention to the spider," "Pay attention to time," and "Pay attention to your intention to push the button." Presumably, readers know how to follow the first instruction (e.g., look at the spider, follow its movements), but how will they respond to the last two? With regard to the second instruction, notice that it does not direct the reader to "Pay attention to the clock." With regard to the last instruction, is the intention to push the button an occurrence that readers can focus their attention on? If the answer is unclear, then consider the following related question: When one writes a sentence (an intentional act), is there an inner urge or feeling for each word, a feeling that could be the focus of the writer's attention?

Assess the effects of the critique on the purposes of the argument. After critiquing the structure of the authors' arguments, one might continue a conceptual analysis with an assessment of the consequences of the critique. For example, on the one hand, the critique of Lau et al.'s (2004) arguments leads us to question the conceptual foundations of the experiment (i.e., the view that intentions are mental or neural representations that efficiently cause behavior and that could be the targets of attention). On the other hand, the critique leaves intact the authors' novel and interesting empirical discovery that the participants' behavior and fMRIs changed reliably with the experimental instructions.

It could be argued that our critique misinterprets the concept of intention, or at least how the authors intended the concept of intention to be used. If so, the source of any misinterpretation is that the authors did not provide a theoretical definition of intention and/or make explicit how the concepts of intention, attention, neural representation, and motor behavior are interrelated. Without some degree of conceptual explicitness, particularly when the argument includes concepts belonging to disparate vocabularies and possessing different grammars (cf. intention and attention, neural firings, motor actions), the door is left open to theoretical confusion (see Bolacchi, 2004). The foregoing aspects of conceptual analysis were embodied in one of Karl Popper's (1959) epistemological theses: "I shall ... adopt a rule not to use undefined concepts as if they were implicitly defined" (p. 75).

Alternative explanations. If Lau et al.'s (2004) interpretation is problematic, then how did the participants reliably comply with the experimental instructions-to attend to the intention to press a button, an intention conceived of as an inner urge or feeling to act-if no such feeling or urge exists qua intention to act? One possible answer is that because the experiment used a within-subjects design in which the participants received separate instructions across conditions (pay attention to the act itself; pay attention to the intention to act) and practiced during one preliminary session, the participants constructed a "script" of how they should behave that was consistent with the instructions and then followed this script by reporting a "moment of intention" slightly before the action. Cognitive psychologist Pylyshyn (2003) made a similar suggestion when discussing whether mental images are

pictures. In an article subtitled "Are There Really Pictures in the Brain?" he wrote,

Nearly all experimental findings cited in support of the picture theory can be more naturally explained by the hypothesis that when asked to imagine something, people ask themselves what it would be like to see it, and they then simulate as many aspects of this staged event as they can and as seem relevant. (p. 113)

Just as mental images may not be pictures, an intention may not be an inner urge or feeling. Instead, the people in Lau et al.'s (2004) experiment may have attempted to make sense of the instructions by imagining what it must be like to have an inner urge before performing an action, and then followed this script by reporting the "moment of intention." Although this account is speculative, it is plausible and testable. Moreover, it illustrates how a conceptual analysis may contribute to the growth of knowledge by checking the conceptual foundations of an argument and inviting new ways to reinterpret important and provocative findings.

### Restoring Conceptual Investigations to Their Rightful Place in Psychology's Scientific Method

What was done in the preceding examples? We did not check the methodological soundness of experiments or the correctness of statistical analyses, for neither of those sets of activities is the domain of conceptual analysis. What we did can be classified into three core activities: (a) examining whether concepts have clear meanings, (b) rendering explicit the grammar according to which concepts are coordinated, and (c) determining whether a set of theoretical statements is consistent. If Cronbach and Meehl's (1955; see also Meehl, 1991, chapters 1 and 10) nomological network comprises nodes (theoretical concepts) that are connected by strands (lawful relationships holding between the concepts), then the first set of activities targets the network's nodes, the second illuminates the strands, and the third assesses their consistency. In what follows, we briefly characterize each set of activities.

The first set consists of identifying a theory's main concepts, characterizing their domains, and understanding their meanings by studying how they are used in specific cases. Thus, when a researcher wonders why we perceive the world around us as erect if our retinal images are upside down, the conceptual analysis identifies a crucial assumption related to the concept of seeing, namely, that we see our retinal images. Similarly, when a researcher explains that generalized or free-floating anxiety is linked to space and time in the same sense that anxiety can be linked to a rat (see Drury, 1973), a conceptual analysis detects a shift in meaning from "Little Albert is afraid of rats" to "Little Albert is afraid of space" and examines what the shift entails (e.g., Is the theory of free-floating anxiety subject to empirical refutation if space and time are conceived as stimuli?). As a rule of thumb, the more open and inferential the set of theoretical concepts or the more unfamiliar their domains, the greater the need for the first set of activities

comprising conceptual analysis (Cronbach & Meehl, 1955).

The second set of activities consists of increasing the explicitness of a theory by examining the grammar of its concepts. The importance of this activity derives from the fact that many theoretical concepts lack explicit definitions and are instead defined by implicit relations with other concepts. The role of conceptual analysis is to render these relations, whatever their source or justification, more explicit so that they can be quantified and tested (see Cronbach & Meehl, 1955; MacCorquodale & Meehl, 1954). As a rule of thumb, the more commonsensical the vocabulary of a theory, the more implicit will be its grammar and therefore the greater the need for the second set of activities comprising conceptual analysis.

Finally, no scientist can anticipate all consequences of his or her presuppositions and assertions. The third set of activities consists of checking the consistency of the theory as it applies to a particular case. The sorts of questions one might ask are as follows: Are there unacknowledged assumptions and steps in an argument? Are the theory's arguments sound? Are the theoretical statements under scrutiny empirically sensitive or necessarily true (see Smedslund, 1995)? As a rule of thumb, the greater the degree of explicitness, articulation, and quantitative formalization of a theory, the more likely the theory is to engage the third set of conceptual analytic activities.

This view of conceptual analysis is related to construct validity (see Cook & Campbell, 1979; Cronbach & Meehl, 1955). To validate a construct, researchers must define, clarify, and explicate the network of meanings attached to the construct and make explicit its assumptions. These steps are prior to experimentation and mathematization because unclear concepts, equivocal propositions, invalid arguments, and incoherent accounts cannot, by definition, achieve any form of validity. Cronbach and Meehl (1955) expressed similar ideas when they reminded us that concepts "may or may not be weighted with surplus meaning" (p. 287), but when they are it is incumbent upon the investigator to make the surplus meaning "sufficiently clear that others can accept or reject it" (p. 291); that is, the meaning must be consensible. Moreover, "when the [nomological] network is very incomplete, having many strands missing entirely and some constructs tied in only by tenuous threads, then the 'implicit definition' of these constructs is disturbingly loose" and "the meaning is underdetermined" (Cronbach & Meehl, 1955, p. 294). Conceptual analysis comprises the set of activities that have the linguistic corpus of theory as their object, specific uses and applications of theory as their context, and the objectification and logical evaluation of the semantic and syntactical patterns of theory as their purpose.

### **Criticisms of Conceptual Analysis**

In this article we have focused on the importance of conceptual/theoretical investigations. And although conceptual analysis is a distinct component of the scientific method, the necessity of which is recognizable in either 17th-century physics or 21st-century psychology, our preceding analyses and arguments may themselves elicit criticism. We anticipate three major ones.

**Conceptual analysis is obvious and obviously a part of psychology's scientific method.** Some may argue that the conceptual problems in the preceding examples were obvious and that, as we have characterized it, conceptual analysis is only sound reasoning. For these reasons, few psychologists deem it necessary to discuss conceptual analysis explicitly, dedicate entire courses to it, or write books and articles about it. Such a view, though, contradicts the long-standing practice of writing about, promoting, and teaching what is critical to experimental psychologists' daily professional practices.

With regard to the obvious, we caution against such claims. In Piaget's well-known number conservation task, a 7-year-old concludes that despite the rearrangement of a row of candies, the number of items *obviously* remains the same because none were added or subtracted. But this same 7-year-old does not remember that two years earlier she concluded that the number of candies *obviously* increased when they were spread out because the row of candies was longer. Conceptual errors, statistical fallacies, and design flaws are more obvious to a good reviewer than a good author (see e.g., Abelson, 1995; Campbell, 2002).

We should also be cautious of claims that conceptual analysis is just sound reasoning. In the broadest sense, the foundations of experimentation and mathematization may also just be sound reasoning, a fact that does not diminish their value. Conceptual analysis is sound reasoning, but it is sound reasoning applied to the verbal output of scientists, to their theories, accounts, models, and concepts. Probing these verbal outputs for their intelligibility is comparable to probing experiments for their methodological soundness or quantitative models for their mathematical correctness.

**Conceptual analysis is a form of verbal tyranny.** Some may say that conceptual analysis is a form of policing people's words and accounts, of repressing the freedom to promote theoretical novelty. Conceptual analysis, in other words, is in conflict with the tolerance necessary for scientific growth. But to say that conceptual analysis stifles theoretical creativity is akin to saying that the rules of logic stifle thinking, the rules of research design stifle experimental creativity, or that the rules of mathematics stifle quantitative creativity.

In a similar vein, we may recognize different "tastes" when it comes to the use of constructs, metaphors, and theories and different tolerances for conceptual and theoretical ambiguity. To paraphrase the proverb, trying to close the door to all sources of ambiguity and error will also leave truth outside. But tolerance for ambiguity and even error does not mean that one should not strive to reduce them. Conceptual analysis helps to detect and reduce ambiguity and error and thereby increase theoretical precision.

**Conceptual analysis is negative activity.** This argument contends that conceptual analysis amounts to destructive criticism, whereas what psychology needs most is constructive theorizing. A related concern is that conceptual analysis tells scientists only what not to do, whereas what scientists need most are guidelines for what to do. Indeed, conceptual analyses can have a negative flavor-not unlike the peer-review process. But it can also be conceived positively in at least two interrelated ways: as part of the selection process of science (Hull, 1988; Popper, 1979) and as part of scientific criticism (e.g., Abelson, 1995). Through mental mutations and recombinations, as it were, scientists engender new hypotheses and theories and then subject them to two broad types of selection. One is based on the empirical adequacy of the scientist's conjectures, and the other on their conceptual clarity, explicitness, and consistency. Observations and experiments, on the one hand, and conceptual analyses, on the other hand, are filters through which all scientific hypotheses, models, and theories must pass. In one of his epistemological theses, Karl Popper (1960/1985) characterized the latter filter as follows:

There is no criterion of truth at our disposal, and this fact supports pessimism. But we do possess criteria which, *if we are lucky*, may allow us to recognize error and falsity. *Clarity* and *distinctness* are not criteria of truth, but such things as *obscurity* or *confusion* may indicate error. Similarly *coherence* cannot establish truth, but *incoherence* and *inconsistency* do establish falsehood. And, when they are recognized, our own errors provide the dim red lights which help us in groping our way out of the darkness of our cave. (p. 55, italics added)

And as with all scientific criticism, we need to offset its perceived, immediate costs with its constructed, long-term gains (see Lyons, 1965; Pearson, 1892/1957, pp. 54–55).

### Conclusions

Ziman (1967) stressed the social dimension of science when he defined it as a search for consensus within a community of rational and competent researchers. But for scientific statements to be consensible, they must reconcile not only facts with facts but also minds with minds. When they fail to do the former, they are empirically false, and science becomes social convention; when they fail to do the latter, they are unintelligible, and science becomes an idiosyncratic representation of reality. The way to reconcile facts with facts and minds with minds is to emphasize all three clusters of activities involved in the scientific method.

Despite its importance for scientific progress, conceptual analysis is acknowledged less than experimentation and mathematization (see the following philosophers and historians of science: Bachelard, 1934/1984; Hacking, 1983; Laudan, 1977; Whewell, 1989; psychologists: Green, 1992; Katzko, 2002; Machado et al., 2000; Slife & Williams, 1995; and biologists: Ghiselin, 1989; Vogel, 1998). The reasons why are undoubtedly many and varied (Machado et al., 2000). We mention one of a historical nature. As heirs of the scientific revolution-seen mostly through the lenses of the Enlightenment-psychologists know the importance of its two major novelties: experimentation and quantification. Naturally, there has been less explicit concern for that aspect of science that existed before the scientific revolution, that aspect present in medieval and classic Greek

science, namely, conceptual analysis. Science has always included the screening of concepts and arguments for clarity and coherence. Archimedes, one of the great classical scientists, discoursed about the lever using the axiom-theorem style of reasoning, in which issues of conceptual clarity and logic take center stage. However, even during the course of the scientific revolution, conceptual analysis remained a part of Galileo's scientific method and even of Francis Bacon's description of method (see Hacking, 1983, chapter 15). It was not abandoned in favor of experimentation and mathematization.

We return to where we started. The scientific method comprises a large set of activities. This set may be partitioned in several different ways, but in the end one should always recognize three subsets or clusters of activities: (a) performing controlled experiments, systematic observations, and correlational studies, (b) framing mathematical or statistical laws, and (c) analyzing concepts and theories. Any partition that yields only two subsets is an impoverished view of method. If our arguments are correct, then psychologists should replace the currently dominant view of the scientific method with one that assigns conceptual analysis its proper weight. This richer view of method would express historian of science William Whewell's (1989) famous dictum that science consists of the colligation of facts and the clarification of concepts.

#### REFERENCES

- Abelson, R. P. (1995). *Statistics as principled argument*. Hillsdale, NJ: Erlbaum.
- Bachelard, G. (1984). *The new scientific spirit* (A. Goldhammer, Trans.). Boston: Beacon Press. (Original work published 1934)
- Bolacchi, G. (2004). On "social sciences" and science. *Behavior and Philosophy*, 32, 465–478.
- Boneau, C. A. (1990). Psychological literacy: A first approximation. American Psychologist, 45, 891–900.
- Broad, W., & Wade, N. (1982). *Betrayers of truth: Fraud and deceit in the halls of science*. New York: Simon & Schuster.
- Burnham, J. C. (1987). How superstition won and science lost: Popularizing science and health in the United States. New Brunswick, NJ: Rutgers University Press.
- Campbell, S. K. (2002). Flaws and fallacies in statistical thinking. New York: Dover.
- Catania, A. C., & Reynolds, G. S. (1968). A quantitative analysis of the responding maintained by interval schedules of reinforcement. *Journal* of the Experimental Analysis of Behavior, 11, 327–383.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design & analysis issues for field settings.* Boston: Houghton Mifflin.
- Crombie, A. C. (1964, May). Early concepts of the senses and the mind. *Scientific American*, 210(5), 108–116.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281–302.
- Drake, S. (1957). *Discoveries and opinions of Galileo*. New York: Anchor Books.
- Drake, S. (1990). *Galileo: Pioneer scientist*. Toronto, Ontario, Canada: University of Toronto Press.
- Drury, M. O'C. (1973). *The danger of words*. London: Routledge & Kegan Paul.
- Galilei, G. (1914). *Two new sciences* (H. Crew & A. Salvio, Trans.). New York: Dover. (Original work published 1638)
- Galilei, G. (1957). The assayer. In S. Drake (Ed.), *Discoveries and opinions of Galileo* (pp. 229–280). New York: Anchor Books. (Original work published 1623)

- Galilei, G. (2001). Dialogue concerning the two chief world systems (S. Drake, Trans.) New York: Modern Library. (Original work published 1632)
- Gallistel, C. R. (1990). *The organization of learning*. Cambridge, MA: MIT Press.
- Ghiselin, M. T. (1989). Intellectual compromise. New York: Paragon House.
- Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review*, 84, 279–325.
- Gibbon, J. (1981). On the form and location of the psychometric bisection function for time. Journal of Mathematical Psychology, 24, 58–87.
- Gower, B. (1997). Scientific method: An historical and philosophical introduction. New York: Routledge.
- Green, C. D. (1992). Is unified positivism the answer to psychology's disunity? American Psychologist, 47, 1057–1058.
- Gregory, R. L. (1996). *Mirrors in the mind*. Oxford, England: Oxford University Press.
- Hacking, I. (1983). *Representing and intervening*. New York: Cambridge University Press.
- Harcum, E. R., & Rosen, E. F. (1993). *The gatekeepers of psychology: Evaluation of peer review by case history*. Westport, CT: Praeger.
- Herrnstein, R., & Boring, E. G. (Eds.). (1965). A source book in the history of psychology. Cambridge, MA: Harvard University Press.
- Hull, D. (1988). Science as a process: An evolutionary account of the social and conceptual development of science. Chicago: University of Chicago Press.
- Katzko, M. W. (2002). The rhetoric of psychological research and the problem of unification in psychology. *American Psychologist*, 57, 262– 270.
- Lau, H. C., Rogers, R. D., Haggard, P., & Passingham, R. E. (2004, February 20). Attention to intention. *Science*, 303, 1208–1210.
- Laudan, L. (1977). Progress and its problems: Toward a theory of scientific growth. Berkeley: University of California Press.
- Leak, T. M., & Gibbon, J. (1995). Simultaneous timing of multiple intervals: Implications for the scalar property. *Journal of Experimental Psychology: Animal Behavior Processes*, 21, 3–19.
- Lyons, J. (1965). A primer of experimental psychology. New York: Harper & Row.
- Machado, A., Lourenço, O., & Silva, F. J. (2000). Facts, concepts, and theories: The shape of psychology's epistemic triangle. *Behavior and Philosophy*, 28, 1–40.
- MacCorquodale, K., & Meehl, P. E. (1954). E. C. Tolman. In W. K. Estes, S. Koch, K. MacCorquodale, P. E. Meehl, C. G. Mueller, Jr., W. N. Schoenfeld, & W. S. Verplank (Eds.), *Modern learning theory: A critical analysis of five examples* (pp. 177–266). New York: Appleton-Century-Crofts.
- Meehl, P. E. (1991). Selected philosophical and methodological papers. Minneapolis: University of Minnesota Press.
- Pearson, K. (1957). *The grammar of science*. New York: Meridian Library. (Original work published 1892)
- Peters, D. P., & Ceci, S. J. (1982). Peer-review practices of psychological journals: The fate of published articles, submitted again. *Behavioral* and Brain Sciences, 5, 187–255.
- Popper, K. (1959). *The logic of scientific theory*. New York: Basic Books.
- Popper, K. (1979). Objective knowledge: An evolutionary approach (rev. ed.). Oxford, England: Clarendon Press.
- Popper, K. (1985). Knowledge without authority. In D. Miller (Ed.). *Popper selections*. Princeton, NJ: Princeton University Press. (Original work published 1960)
- Pylyshyn, Z. (2003). Return of the mental image: Are there really pictures in the brain? *Trends in Cognitive Sciences*, 7, 113–118.
- Quine, W. V. (1936). Truth by convention. In O. H. Lee (Ed.), *Philosophical essays for Alfred North Whitehead* (pp. 90–124). New York: Longmans, Green.
- Slife, B. D., & Williams, R. N. (1995). What's behind the research? Discovering hidden assumptions in the behavioral sciences. London: Sage.
- Smedslund, J. (1995). Psychologic: Common sense and the pseudoempirical. In J. A. Smith, R. Harré, & L. Van Langenhove (Eds.), *Rethinking psychology* (pp. 196–206). London: Sage.
- Stanovich, K. E. (1998). How to think straight about psychology (5th ed). New York: HarperCollins.

- Uller, C., Hauser, M., & Carey, S. (2001). Spontaneous representation of number in Cotton-Top Tamarins. *Journal of Comparative Psychology*, 115, 248–257.
- Vogel, S. (1998). Academically correct biology. *American Scientist*, 86, 504–506.
- Wade, N. J. (1998). A natural history of vision. Cambridge, MA: MIT Press.
- West, R. E., & Young, R. J. (2002). Do domestic dogs show any evidence of being able to count? *Animal Cognition*, 5, 183–186.
- Whewell, W. (1989). *Theory of the scientific method* (R. E. Butts, Ed.). Indianapolis, IN: Hackett.
- Wynn, K. (1992). Addition and subtraction by human infants. *Nature*, 358, 749-750.
- Wynn, K. (1995). Infants possess a system of numerical knowledge. New Directions in Psychological Science, 6, 172–177.
- Ziman, J. (1967). *Public knowledge*. Cambridge, England: Cambridge University Press.