Scriven, M. (2012). The logic of valuing. In G. Julnes (Ed.), *Promoting valuation in the public interest: Informing policies for judging value in evaluation. New Directions for Evaluation, 133,* 17–28.

2

The Logic of Valuing

Michael Scriven

Abstract

This chapter outlines the logical infrastructure that makes it possible to claim that one can validate values, both at a general and a context-specific level, other than by direct deduction from other value premises. To make the argument, the author distinguishes between the logic of valuing and the logic of evaluation, the former being the primary focus of the argument, and analyzes the invalidity of the long-standing value-free doctrine in social sciences. The author discusses several ways in which we can establish factual as well as evaluative premises by observation, inference, or definition, that make it possible to infer beyond reasonable doubt to evaluative conclusions. © Wiley Periodicals, Inc., and the American Evaluation Association.

Hume is frequently quoted as having said, "Reason is the slave of the passions." What he actually said was more severe: "Reason is *and ought to be* the slave of the passions *and never pretend to any other office than to serve and obey them*" (italics added). Hume was talking about human passions in the sense of all that we desire or value, not necessarily passionately, and it is clear that as a matter of fact our desires or values include many that can be changed by reason, such as the desire for an electronic gadget, bottled drink, drug, or food, that turns out to have an extremely dangerous design flaw or side effect. Indeed, our desires or values also include *passionate* desires that can and should—at least for most people—be changed by reason, such as the desire for cigarettes or methamphetamine.

So there is a place for reasoning about values, and we even do it, some of the time, and rather obviously should do it more of the time. Hence—if we take him literally—Hume was wrong, both descriptively and prescriptively. A more plausible aphorism in our time of atomic bombs and atomic power plants, biological warfare, global warming, and vanishing natural food and fuel resources, might be: "Reason is often the slave of the passions today *but tomorrow only comes if it rules them.*"

In the profession of evaluation, reason appears to exert some sway over the passions, because our careful investigations to support an evaluative conclusion often appear to weigh one valued outcome against another on the basis of certain factual or purely logical considerations, for example, differential cost or feasibility or consistency. Whether or not there is always an evaluative premise involved, on which Hume's position depends, this is a radical transformation of attitude from the days of the more extreme value-free doctrine about science—most of the 20th century—when reasoning about values and value judgments was held to be impossible within the domain of science; hence program evaluation and other branches of the discipline of evaluation were disbarred from consideration as scientific activities. Although this skeptical conclusion is now widely rejected by social scientists—for example, the social scientists who are authors of a leading text on evaluation define evaluation as applied social science—it is still common, perhaps even dominant, as an underlying prejudice in those disciplines, as we can see from recent anthologies on applied social science in which the term *evaluation* and examples of its practice are entirely absent. What I want to do here is to outline the logical infrastructure that makes it possible to claim that one can validate values, both at a general and a context-specific level, other than by direct deduction from other value premises. To do this thoroughly would take more than the limited number of words than I have, but I can at least indicate the lines of argument that I think are adequate for these purposes.

Preliminaries

We must begin with an understanding of two matters. First, we must understand the distinction between the logic of valuing and the logic of evaluation. The former is our main topic and Hume's topic—the logic of supporting or refuting our commitment to valuing, that is, liking, admiring, or appreciating certain things or tastes or attitudes to which we have become attached. The latter is the logic of the process of applying those likings and valuings to determining the value of *other* things whose value is, for the moment at least, not yet determined or not determined with sufficient certainty for some purpose. These processes are related and overlap in complex ways that we intuit better than we have so far explained, but that should become a little clearer here, where we will use some of the logic of evaluation to get to our conclusions about valuing.

Second, to get a grip on either of these logics, we need to understand the reasons the value-free doctrine was attractive to both positivists and social scientists in the early decades of the 20th century, and remains so for many social scientists today. The logics of valuing and evaluation exist only because these reasons for denying the legitimacy of both are invalid, and they can be understood and accepted only if we understand the reasons for those errors. I think there were eight main errors involved in the value-free doctrine, of which the first seven were based on an oversimplified philosophy of science inspired by the success of the existing physical sciences, and formulated by the positivists. I'll list them and then indicate why they do not show evaluative claims to be less scientific than the usual scientific claims made by the social scientists proclaiming or still attracted by the value-free doctrine. This will show that the supposed arguments against any logic of valuing or evaluation are flawed, which makes our enterprise here at least possible

First, social scientists thought that the typical value claim was an expression of personal taste, hence entirely subjective in the sense of being idiosyncratic and hence not generalizable—hence not a scientific claim, in their view of science as a discipline primarily interested in general laws and truths, often expressed by saying science was nomothetic rather than idiographic like history. They would often say, "What's good for one person, or valued by one person, isn't good for or valued by another—but that's not the way scientific claims hold true—they're true for all (suitably trained) observers." Their ideal sciences were the great nomothetic sciences, physics and chemistry, but today we are well aware of the virtues of largely idiographic subjects like geology, planetology, forensics, single-subject comparative and clinical psychology, and epidemiology. The positivists thought that idiographic studies were largely prescientific, or merely confirmatory, or anecdotal. And, of course, much of program and personnel evaluation consists of the study of single cases. But this reason is no longer plausible, because single-subject science and idiographic sciences are acceptable.

The second mistake was a kind of reverse edge on the first one, but logically distinct: it was the view that value judgments were also subjective in a second sense, meaning that value judgments could not be confirmed by others (as well as not being true of others). This error was based on a simplistic view of confirmation as direct sensory observation of external phenomena. In fact, other people can confirm claimed matters of taste indirectly by observations of the claimant's behavior, and evidence of his or her veracity in other reports. Both of these errors could have been avoided if they had looked carefully at their own use of value judgments, e.g., in evaluating the work of their students and peers, or in evaluating the experimental designs or scientific instruments they used at work, since those evaluations were supportable by reasons that others could check.

Apologists for the value-free doctrine sometimes say, "Well, what was really meant by *value judgments* was just judgments about ethical and

cultural values, which of course *are* like matters of taste and can't be justified by reasons. It's a bad defense because (a) all their arguments were couched in logically general terms, that is, about all value claims, not a subclass; (b) even if there were not good reasons in the abstract for this subset of values, there are plenty of contexts in which there are good reasons for treating them as acceptable premises; (c) context apart, there are plenty of good reasons for or against many, perhaps most, cultural and ethical values (more on this below). The proper defense against illicit smuggling of political values into science is simple; demonstrate their irrelevance if you can. Just don't invent garbage to defend against garbage.

Third, positivists and social scientists had no conception of the way in which the meaning of valuable concepts could be fundamentally context dependent, and still be entirely scientific. Relativity theory had not yet reached the acceptance status of proving that, for example, velocity was fundamentally context dependent, and they didn't think the vast range of context-dependent terms in ordinary language—terms like *tall* or *cheap* or *very large*—was an acceptable type for science. That was an error, and again, scientists' own practice demonstrated that *good* is a highly context-dependent concept, and nonetheless entirely capable of objective use. Their own defense of claims like "General relativity provides a very *good* explanation of some solar eclipse phenomena" or "Einstein was a very *good* theoretical physicist" vitiates the contrary view.

Fourth—the second edge of their objection to terms like *tall* and *cheap*—positivists thought that precision was crucial (or at least very important) for scientific concepts, again not noticing their own use of terms like *original* or f*undamental*, or *analogous*, which have never been quantified. (Today we would add *significant* where the failure of quantification is now enshrined as a great learning opportunity.)

Fifth, a cousin of the fourth mistake, they thought that approximations had no place in laws of nature, although they came around to accepting statistical laws in quantum physics. This was unfortunate, because virtually all the laws positivists used in their examples—the propositions normally called laws in physics and chemistry texts—are in fact only approximately correct, and not very close to the truth across vast ranges of their variables (Scriven, 1961). This rendered invalid one of the objections to the way in which ethical (and other evaluative) generalizations such as "Thou shalt not kill" are maintained in spite of known counterexamples. (The view that scientific generalizations had no exceptions was Karl Popper's favorite hobbyhorse.)

Sixth, positivists had an extremely narrow-minded view of *observation*, and it remains perhaps their most strongly felt as well as their most egregious legacy today. In particular, they would not accept three types of observation that careful commonsense epistemology finds useful and valid: (a) observations of one's own inner states—*introspection*, as in noting the height or depth of one's despair or joy; (b) observation of causation as in

seeing someone knock over a cup; and (c) observation of value or merit, as in observing the excellence of a maneuver by a skater, gymnast, calligrapher, mathematician, or soccer player. Obviously one can make mistakes about all of these types of observation, and lie about some of them; but that's true of all observation, which for the positivists was the epistemological foundation of all science: We just have to know how to cross-check, triangulate, and look at the observer's training and track record.

The seventh reason was a naïve concept of *justification*: the idea that for a belief to be justified, it had to be strictly inferred from other beliefs that could be independently established, for example, by observation or logico-mathematical proof. Hence evaluative claims could never be justified, because they could only be inferred from other evaluative claims, which could themselves never be independently justified (since they could never be directly observed, or established as required explanations of what could be directly observed). Of course, once one adopts the commonsense view of what can be observed, we now have to accept factual, observationally verified, evaluative premises, from which we can infer and hence justify some evaluative conclusions in exactly the way the positivists thought impossible.

The eighth factor, a cause but not an explicit belief, and only operative for some of the skeptics about evaluation, was *valuephobia*—the irrational fear of, and extreme opposition to, evaluation that almost every evaluator has encountered in their professional practice. The idea of cutting the wolf loose—legitimating the process of systematic evaluation—understandably made the early social scientists nervous because they were insecure about the credentials of their disciplines.[1] Even today the practice of evaluation, or quality control, is seen as a major threat by many in academia as well as in business or government. Nervousness about evaluation is sensible, criticism of specific approaches to it should be encouraged, but blind opposition is irrational. The sciences are only acceptable as sources of truth as long as their processes of internal verification are valid, that is, they are completely dependent on evaluation for validity. But human nature and nurture make some people supersensitive to criticism, even possible criticism, beyond the point of invading their own survival orientation. It's a phobic reaction.

Probative Proposal

Now let's turn to a more constructive approach and look in more detail at just how we can provide disciplinary and indeed scientific credentials for these imprecise, context-dependent, approximative concepts that are at the heart of evaluation as well as everyday communication. It's good to recall just how this epiphany worked when statistical propositions gradually acquired scientific status, after a beginning during which they were sneered at, even by many mathematicians, as impostors. In just that kind of way, we must extend our logic to handle the new imprecision by (a) specifying the

context (e.g., by doing careful needs assessment) instead of throwing up our hands at the fact that context is almost always a key element of the meaning of the very important vocabulary of evaluation; (b) making allowances for inexactitude in our concepts, just as we made allowances for statistical truths, instead of dismissing all vague terms as unacceptable; (c) accepting laws that are only approximations if they bring order out of chaos, that is, exhibit a far-reaching pattern that is close to the truth in many cases of interest, and probably indicative of an underlying microexplanation; and (d) expanding our notion of what is observationally verifiable to include some carefully selected and checked, causal, introspective, and evaluative claims. Spelling out the details of how to make these four adjustments would require more space than we have here, but what are perhaps the two most important moves will be outlined. (e) More controversially, we will need to adopt an extension or modification of what is commonly thought of as the set of permissible inference rules on which all valid reasoning depends. The pinch that the limitation to deduction and statistical inference imposes has long been felt, as one can see immediately from the list of names for some kind of a Third Way that has been proposed.[2] I use the term *probative inference* for my candidate, lifted from the vocabulary of jurisprudence and perhaps slightly extended. It *is* covered by the definition of probative as a legal term in the *Oxford Dictionary of American English*, which is "having the quality or function of proving or demonstrating something," except that I use it by contrast with, rather than as including deduction and probabilistic inference. To place it in the present context, let us go back to the basic logical question that faces any attempt to analyze the logic of evaluation, which is, how can we validate/prove/establish evaluative claims?

The simplest way, already mentioned, is by direct observation, and we'll return to that in a moment. But a more traditional way—in principle—is by providing an evaluative premise that, in conjunction with our empirical premises, makes it possible to deduce the desired evaluative conclusion. (In the technical vocabulary of logic, the premises are then said to entail the conclusion.) The good news for program evaluators is that most evaluations we are commissioned to do only require rather vague evaluative premises that are easy to support, e.g., "The 660,000 tent people in Haiti (as of August 2011) still badly need potable water and more food." Surely it is easy to deduce that the project we are evaluating (the evaluand), which does in fact supply nutritious food and potable water, is valuable. The bad news is that this isn't a deduction, because, among other potential problems, the project may have bad side effects that wipe out the benefits. Well, can't we just check the side effects and then, all being well, add that premise to the others? Unfortunately, checking for all possible side effects is not like checking an integer to see if it's a prime number; it's not a finite task. However good an evaluator you are, and however many possible side effects you have checked for, another adverse one can still show up, years later if not

immediately. Hence the deduction requires a somewhat speculative premise, without which it is impossible. This example is a really simple one, but even in such a case, getting clear about the evaluative premises it requires is quite valuable, because it underlines that the search for side effects is almost always a crucial part of good evaluation. Given that one rarely finds any sections about side-effect searching in current evaluation reports or textbooks, this is a needed reminder that due diligence requires more attention to them. And that's not the only problem; there may be improprieties, perhaps extremely unethical practices, in the way the water and food is distributed. The benefits may save someone bearing cholera or malice who will kill more people than are saved—not a side effect in the usual sense, but a malign idiosyncrasy. And so on. All that will save the deduction is an illicit circular premise—one that says, "there are no other flaws in the premises—the deduction is valid." Evaluative inference is hazardous if deduction is your ideal. As is much scientific inference.

Nor is a statistical inference possible, meaning by this a quantitative claim. Of course, the more investigation one does, without finding seriously negative side effects, and the more competent at that task one is, the more probable it becomes, in a qualitative sense, that the present project is meritorious. But I think it's advisable not to count that as *statistical* or even *probabilistic inference*, terms that, today, suggest some quantitative scale is being applied or approximated. I am going to take that stand on those terms and demarcate another type of qualitative inference, however, as *probative inference*: It is an inference to a conclusion that has been established, so the utterer claims and is prepared to support, as beyond reasonable doubt. (I am not sure of the extent to which this concept overlaps with defeasible or rebuttable or adbuctive inference, as these terms have been used so far.) This is not inference to a conclusion that is beyond the possibility of error, given the truth of the premises; that would be deduction. But it goes beyond mere probability and it makes a claim of defensibility against attack that goes beyond prima facie status: It issues a kind of promissory note for justified confidence, meaning that further defenses are available if challenged. Under cross examination in a court of law, the expert witness rejects (or should reject), defense counsel's aggressive suggestion that "You just mean 'it's probable, in your view.'" The correct response is: "No, it's not just my view and it's not just probable: it's my professional judgment that the evidence establishes this conclusion beyond any legitimate doubt by an objective, reasonable, and competent expert." Of course, this view is often challenged by other experts produced by the defense—it is indeed rebuttable—and their relative expertise then becomes the focus of attention. But reputation is not enough in the eventual court of scientific inquiry; there must be a specific alternative explanation provided in order to make the challenge credible, and the utterer of the probative claim will have the chance to rebut such a challenge, as the promissory commitment indicates. Absent a

successful challenge, we have an inference to the best explanation and good cases of that are good enough to bet one's reputation on—and sometimes, not only in the medical field, to bet one's life on. Their conclusions are not just rebuttable but assertible, and not just assertible but assertible as beyond reasonable doubt.

Moreover, probative inference goes far beyond the domain of explanations: exactly the same epistemology is involved in what we might call inference to the best classification, which we see not only in taxonomy but in the less formal processes of bird-watching, and hieroglyph or bad handwriting interpretation, and even in face and pattern recognition. Indeed, it's clear that it covers inference to the best description. The eliminative and comparative process that goes on in defending a description in a journal, a newspaper, or a scientific report is just like the one that goes on in inference to the best explanation, and both are probative inferences.

The bottom line is that what we have here has essentially the same status as we accord observations, the epistemic gold standard. Observations are also of course fallible, rebuttable, but nevertheless far beyond prima facie truth or mere probability. Although of course a probative conclusion is very highly probable, that's not all it is. It is clear these domains—observation, description, causation, explanation, and evaluation—overlap in their processes of inferential support, and in their eventual certainty status, despite protestations to the contrary. Witnesses do indeed, sometimes, see the defendant shoot the victim dead right in front of them, that is do see, and report as seen, a causal claim: they do not infer to that (Hume was wrong on this, too). Of course, some causal claims *are* inferences but some of these are probative inferences, every bit as beyond reasonable doubt as most observations. For example, when the autopsy reveals only a hole in the forehead and a bullet in the brain—nothing else wrong with the decedent—then the inference to the cause of death is usually probative. This is the informal logic of ordinary inferences and the logic of evaluative inferences, which are part of the folk logic of everyday life at home and in the market and office; and it's not inaccurate talk, it's fully functional and as accurate as is appropriate.

Although one's first reaction to the positivist tightening up of ordinary language is sympathetic, a more sophisticated reaction is a subtler interpretation of that language. One might even come to say that "the whole of science is nothing more than a refinement of everyday thinking. It is for this reason that the critical thinking of the physicist cannot possibly be restricted to the examination of concepts from his own specific field. He cannot proceed without considering critically a much more difficult problem, the problem of analyzing the nature of everyday thinking." Those are the words, spoken in 1936, of a highly credible person who personally knew and was at first much attracted to the positivist position, but came to see beyond it (Einstein, 1936, p. 349).

Is probative inference always the best we can do in supporting evaluative conclusions? John Searle has argued persuasively that there are some

cases when one can actually deduce evaluative conclusions from factual premises, for example, when someone borrows money from a friend and says "I owe you fifty dollars" as he takes it—this being a factual, observational, premise—then they really do have an obligation to repay it, and *that* claim is an evaluative deduction from the factual one. These are rare cases, and have not been accepted as incontrovertible cases of deduction—but they are at least cases of probative inference. And there's one more category of probative inferences that is of great importance and very common in evaluation, an explication of one type of inference to the best classification or identification. It is a subcategory of what I call *criterial inference* (Scriven, 1959).

A classical definition is a statement of meaning equivalence between two word strings: (a) the definiendum (the term being defined) and (b) the definiens (the terms making up the definition). Except for mathematical and a few technical terms, and neologisms until they catch on, there are very few valid classical definitions of terms in common use, either in scholarly or in everyday contexts. Nearly all the definitions in a regular dictionary are merely listings of a few leading criteria, that is, properties that are either part of the meaning but not a necessary part of the meaning; or just indicators, that is, properties that are commonly coexistent with the definiendum. And the total list of the terms in the definition does not add to all of the meaning. The term *apple*, for example, is defined in the Oxford American as "the round fruit of a tree of the rose family, which typically has thin red or green skin and crisp flesh." That includes rose hips on a tree rose, which aren't apples, and excludes baked apples, which are. Golden Delicious, a common variety, are yellow, but typically a definition allows for a few exceptions, thereby conceding a lack of general validity. We can improve these definitions by adding more criteria, for example, edibility when ripe, seeds and a core, stems, and so on; but we lose concision and hence convenience, for example, in speed of comprehension, or of lookup (because of increased dictionary size). Criteria are part of the meaning because it's a sample of these that we mention when explaining the meaning of a term to a learner of the language, young or adult. But there are many of them for most common descriptive terms, and few people can even list them all, although they may recognize them as part of the meaning fairly reliably; and they fade away into mere indicators in a gradual and debatable way. If the term X has the indicators Ci, we can express this situation as X = C1 + C2 + C3 + . . . If this were a classical definition, we could deduce the presence of each of the C's from the presence of X, and the presence of X from the presence of all of the C's. But with criterial definitions, we are considerably limited. We can deduce from X to the conclusion that *some* (unspecified) C's are present, and sometimes we can roughly rank or at least group some of the leading C's in order of the likelihood that they are present in a given context. The winners in this ranking can often be *probatively* but not deductively inferred as present, if X is present. And we can often say that if a specified subset of

C's is present, sometimes the ones in the dictionary, then X is probatively present, and vice versa; those are two important types of criterial inference.

So the claim here is that it is useful to identify a third type of sure inference, along with deduction and very highly probable/statistical inference. The new type is essentially different because it is a pragmatic/rhetorical concept; its meaning is rooted in a potential dialog of challenge and defense that has context-dependent rules. Probative inference has a number of important subtypes: (1) legal reasoning in a criminal court (when the charge is merely a misdemeanor, a lesser standard of proof applies—balance of evidence). This subtype is highly stylized, with explicit rules governing the dialog. (2) Inferences from most laws of nature (or models of phenomena) to specific predictions or explanations—these can't be deductions because the laws are only approximations; (3) inferences to a classification, for example, by a forensic pathologist or birdwatcher; (4) inferences to the best explanation; (5) inferences to evaluative conclusions from some evaluative generalizations (e.g., "Tent people in Haiti badly need food and potable water"); (6) inferences based on very strong analogies or theories, for example, to other minds and other planets; and (7) the fundamental process of induction to the future, which may be another example, but its candidacy would need further support.

Many of the informal logic/critical thinking/argumentation experts would be sympathetic to much of what's just been said. But now we come to something new; it provides us with the eighth strand in a logic of valuing, making up a logic that is rather different from the usual logic of science as represented by philosophers of science even today. (8) When you learn the meaning of a word (call it X) in the usual way—that is, someone tells you, or gives you some examples of what it means (i.e., an ostensive definition), or if you look it up in a dictionary—you have usually learned to associate its meaning with a set of criteria, Ci, with some sense of the relative weight of these. But now I want to suggest that it is a *separate* task, often a longer one, to learn the meaning of good X and/or bad X. For example, it is easy to learn what a mountain bicycle is, or what a word processing program is, but that doesn't tell you the criteria for a good one (though of course you know that set will include the first set). For that you need an expert and/or a research budget plus good product evaluation skills. You don't need all the fine details in order to do serious product evaluation, but you do need at least the section headings in a validated checklist of criteria of merit. (See, for example the checklist of criteria of merit in my professional-level book on evaluating word processors which runs over 35 pages; the section headings would take only four or five pages and are adequate criteria of merit (Scriven, 1983).) Now even a product that scored well on all those criteria could conceivably be seriously flawed; no author is infallible, and technologies change. But anyone who knows the field, even as a consumer rather than a designer, would be likely to agree that a good score

on nearly all these dimensions of merit would probatively imply that the product was a very good one.

If the criteria of merit can be defined or explicated nonevaluatively then that inference is an inference from merely factual premises[3] to evaluative conclusions, using only definitional rules for legitimation. Many of these criteria of merit can be so defined; for example, we usually define reliability of office machinery in terms of mean time between failure (MTBF), and we can say, as an example, that an MTBF of 2 years (for failures that take more than 2 hours for the tech service people to fix) rates as good to very good. Because this is part of the meaning of *good word processor*—just as *sweet tasting* is part of the meaning of *good peach* (for out-of-hand eating)—it's not an extra evaluative premise.

Some of the criteria of merit, in real-world checklists, are normally expressed in evaluative language, but that's not important as long as those criteria, in turn, can be defined criterially in nonevaluative language. (This parallels the process of validating theories by appealing to observations; in fact, we often do it by appealing to claims that have a theoretical basis, too, but that's acceptable if they in turn can be validated without assuming the theory at risk.)

To conclude: I think that the preceding material indicates several ways in which we can establish factual as well as evaluative premises by observation, inference, or definition, that make it possible to infer beyond reasonable doubt to evaluative conclusions. Such reasoned inferences establish the truth of conclusions about valuing or disvaluing many things—in certain contexts, and to varying degrees—including passions for addictive drugs, hatred of ethnic groups, and the fundamental ethical attitude of valuing the prima facie rights of all humans (Scriven, 1966). So Hume was wrong, but more importantly, so were many of the reasons we have accepted for believing him.

Notes

1. Indeed, the value-free doctrine appears to have originated in Max Weber's cautions to early sociologists against scientific investigation of government programs, based on his concern about government funding reprisals. That was understandable caution, even if not exactly a testimonial for the search-for-truth value that science is supposed to embody; but the subsequent move to rationalize it with an ill-founded doctrine of value-free science was surely in part valuephobic, as we can see from the extent to which it involved ignoring their own evaluative and scientific practice.
2. New trivalent logics have been called nonmonotonic, default, autoepistemic, paraconsistent, and relevant; the proposed third type of reasoning has been called plausible, prima facie, presumptive, defeasible, analogical (John Wisdom), conductive (Wellman), deduction (Sherlock Holmes), rebuttable (Toulmin), argumentative (Perelman), and the somewhat mysterious abduction (Peirce). (Thanks to Tony Blair for helpful suggestions for inclusion here.) The term *multivalued logic* has been used for a related but not identical amplification of classical logic—the inclusion of more than the two truth values of true and false.

3. What are commonly called *factually true premises*, also known as *matters of fact*, or—speaking slightly more technically—*empirically true premises*, are not, speaking precisely, always nonevaluative, because many evaluative propositions are clearly factual or—at least in one sense—empirical; for example, "Einstein was a great physicist," or "Scoring more than 150 on the usual I.Q. tests, without cheating, means at least that the respondent is an unusually good problem-solver." But here we'll occasionally use the term *factual* (or merely, or purely, or simply factual), or *empirical* (etc.) for brevity, as they are intended colloquially, because they are colloquially meant to convey being nonevaluative. (The false dichotomy of facts vs. values is a relic from the positivists.)

References

Einstein, A. (1936). Physics and reality. *The Journal of the Franklin Institute, 221*(3), 349–382.

Scriven, M. (1959). The logic of criteria. *Journal of Philosophy, 56,* 857–868.

Scriven, M. (1961). The key property of physical laws: Inaccuracy. In H. Feigl & G. Maxwell (Eds.), *Current issues in the philosophy of science* (pp. 91–101). New York, NY: Holt.

Scriven, M. (1966). *Primary philosophy.* New York, NY: McGraw-Hill.

Scriven, M. (1983). *Word magic: Evaluating and selecting word processing.* Belmont, CA: Wadsworth.

MICHAEL SCRIVEN is a professor of psychology at Claremont Graduate University, a senior research associate at the Evaluation Center, Western Michigan University, and director of institutional research at Palo Alto University. He has served as president of the American Educational Research Association and the American Evaluation Association.