

On the Purposes Served by Psychological Research and Its Critics

Mark Schaller

UNIVERSITY OF BRITISH COLUMBIA

Christian S. Crandall

UNIVERSITY OF KANSAS

ABSTRACT. We concur with Wallach and Wallach's (1998) subjective assessment that much psychological research contributes little to our corpus of knowledge, but we dispute their analysis of the causes of this problem. A critical assessment of their analysis reveals it to be (a) logically flawed, (b) irrelevant to hypotheses concerning psychological processes, and (c) potentially injurious to the processes through which creative scientific hypotheses are developed. The Wallachs' article may serve a valuable purpose—but only if read very critically.

KEY WORDS: criticism, deduction, falsifiability, hypotheses, tautologies

Edwin Denby once said that 'it is not the critic's historic function to have the right opinions, but to have interesting ones'. Denby was a dance critic, and was speaking about performance of art, not science. There are some differences between art and science. The merit of scientific practices is measured not merely against aesthetic criteria, but against epistemological criteria as well. Similarly, criticism of scientific practices may also be judged against epistemological criteria: To what extent are these criticisms logically sound? To what extent might these criticisms contribute to the accumulation of scientific knowledge? It is with these two questions in mind that we criticize the critique offered by Wallach and Wallach (1994, 1998).

We have done so once before (Schaller, Crandall, Stangor, & Neuberg, 1995). On the basis of our earlier audition as critics-of-the-critics, Wallach and Wallach have miscast us instead in the role of 'mainstream' reactionaries who 'champion the testing of near-tautologies' (p. 191). This miscasting misses the fundamental points of our criticism. In fact, we share the Wallachs' subjective feeling that 'much customary psychological research serves little purpose' (p. 192). We too would prefer to see more research that attempts boldly to break the boundaries of existing knowledge. We are not alone in sharing this impression; even the most prolific 'mainstream'

researchers in experimental psychology offer the opinion that 'journal articles just aren't interesting anymore' (Higgins, 1992, p. 489), and that the hypotheses being tested are 'safe yet boring' (Wegner, 1992, p. 506). Everyone agrees: a lot of experiments in psychology are pretty darn dull.

But we disagree with the Wallachs' analysis of this problem. They claim that the problem occurs because 'researchers frequently attempt to test hypotheses which are, or are derivable from, inherently unfalsifiable propositions' (p. 183). These propositions they call 'near-tautologies'. A careful examination of the Wallachs' arguments reveals their logical analysis to be unsound. There is no evidence for their contention that psychologists frequently test unfalsifiable hypotheses. If that was the only issue, however, we wouldn't really care. What concerns us much more is that, if accepted uncritically, the Wallachs' inaccurate analysis of the problem may not only fail to address the problem, but may actually contribute to it instead.

The Wallachs' Logical Analysis is Flawed

Words can have multiple meanings. The word 'rug' may refer to a woven mat intended to cover a floor, or to a woven mat intended to cover a balding scalp. People walk around on their 'rugs'; people affix 'rugs' to their heads; but if we tried to convince you, therefore, that people walk around on their balding heads, we hope you would laugh.

The same logical flaw underlies the Wallachs' claim that 'researchers frequently attempt to test hypotheses which are, or are derivable from, inherently unfalsifiable propositions' (p. 183). They assert that hypotheses derivable from 'near-tautologies' are unfalsifiable. They assert that hypotheses derivable from 'near-tautologies' are frequently tested. It is no surprise that they reach the conclusion that they do. But it turns out that a 'near-tautology', like a 'rug', can mean different things, and the meanings are not consistent across the Wallachs' two assertions.

The Wallachs' intended definition of 'near-tautology' (which we call the 'strong' form) is illustrated by the example 'Interest in a target's traits is likely to increase attention to trait-relevant information'. At a conceptual level, this proposition is not a tautology—neither interest nor attention is conceptually defined in terms of the other. At an operational level, however, they may indeed be defined in a circular fashion. Any psychological hypothesis that fits this structure might well be unfalsifiable—if not in a truly logical sense, then certainly in a psychological sense. No scientist would dismiss such a hypothesis merely because sloppy data collection failed to support it. Experiments testing hypotheses of this sort would certainly contribute little to our corpus of knowledge.

A second definition (a 'weak' definition) of 'near-tautology' is implied by virtually every other example provided by Wallach and Wallach in their two

articles. Consider the following proposition that the Wallachs assert to be 'near-tautological': 'Increased attention to trait-relevant information is likely to increase accuracy of trait ratings' (p. 185). Compare closely this weak-form 'near-tautology' with the strong form above; the weak form is a different beast altogether. Attention and accuracy are neither conceptually defined in terms of one another, nor are they ever operationally defined in a circular manner. Operational definitions of attention are *not* embedded in conceptualizations of accuracy; operational definitions of accuracy are *not* embedded in theories of attention. The statements are not definitionally circular, but rather express a causal relation that may seem, to some observers, obviously correct. (Confidence in the relation may result from knowledge of previous empirical demonstrations, from advocacy of a theory that predicts such a relation, or simply from intuition.)

In contrast to strong-form 'near-tautologies', propositions that fit this weak-form 'near-tautological' structure are entirely falsifiable. As we discussed at length in a previous paper (Schaller et al., 1995), any claim otherwise confuses logic with psychology, and confuses the actions of individuals with beliefs of collectives. A single scientist might retain belief in the veracity of a hypothesis even in the face of uncooperative data, but that reflects nothing about the logical status of the proposition; it reflects instead the psychological prejudices of that individual. Prejudices vary. While one scientist may cling to an unsupported hypothesis, another scientist rejects the same hypothesis on the basis of the same data. Moreover, ostensible obviousness is a notoriously unreliable indicator of truth. Many well-accepted theories, self-evident hypotheses or intuitively obvious propositions turn out to be wrong, and become recognized as such. These hypotheses are not only logically falsifiable, but become falsified.

So, according to the definitions implied by the Wallachs' analysis, 'near-tautologies' come in not one flavor but two. They may look similar on the surface, but they are as different as chalk and cheese. In demonstrating that 'near-tautologies' are unfalsifiable, the Wallachs have in mind the 'strong' definition. In demonstrating that 'near-tautologies' are commonly tested, the studies they review reveal only propositions that meet the 'weak' definition. Is there any evidence that strong-form 'near-tautologies' are commonly tested, supported, submitted and reported in the psychology journals? No. Not a single one of the actual hypothesis tests deconstructed by the Wallachs fits this structure.

Of course, the Wallachs do not claim that psychologists put 'near-tautologies' explicitly to test; rather, they claim that the hypotheses usually tested by psychologists are logically deducible from 'near-tautologies' and so are themselves unfalsifiable. This subtler argument is based on the same faulty logic, and it too turns out to be unsound.

Any experiments testing hypotheses derived entirely from strong-form 'near-tautologies' are unlikely to be particularly informative. Although one

might justifiably quibble with the Wallachs' characterization of such hypotheses as 'inherently unfalsifiable' (p. 183), they are surely right in claiming that tests of these hypotheses would serve little purpose. But again, there is no evidence that hypotheses derived entirely from strong-form 'near-tautologies' are commonly tested, or ever reported in the psychology journals. Every one of the logical deductions reported by the Wallachs in their articles depends entirely or primarily on weak-form 'near-tautologies'.

Hypotheses derived from weak-form 'near-tautologies' fit the following logical template: previous research, theory and intuition imply (with great confidence) that variable A predicts B, B predicts C, and C predicts D; hence it is hypothesized that A predicts D. The underlying causal logic is certainly compelling, but we have seen that weak-form 'near-tautological' propositions are falsifiable. It turns out that hypotheses deduced from a chain of such propositions are equally falsifiable, and are even more likely to be falsified. When an argument is based on multiple, connected elements, the confidence ascribed to that argument cannot be greater than the confidence ascribed to the least probable element, and is usually much much lower. The statement 'If it is raining on my house right now, it will be raining on my house one minute from now' is accurate to a very high degree of certainty—but it is not quite 100 percent certain. The same logic holds for the statement 'If it is raining on my house one minute from now, it will be raining on my house two minutes from now', and for every other statement about the likelihood of rain on any two consecutive minutes. String a bunch of these statements together, and it is easy to 'logically deduce' that 'If it is raining on my house right now, it will be raining on my house a week from Friday'. And yet, unless your house happens to be on Venus (or in Vancouver), the deduced hypothesis is unlikely to be held with confidence; it will be easily abandoned if, on the appointed day, your house is bathed in sunshine. So too it is with psychological hypotheses derived entirely from weak-form 'near-tautologies'.

The same holds for hypotheses deducible from a mixture of strong- and weak-form 'near-tautologies'; their logical status is no different from hypotheses derived entirely from weak-form 'near-tautologies'. If *any* of the underlying propositions deviates from tautology, then the hypothesis itself is accordingly worthwhile to put to test. Return to the rain example and add the truly tautological statement that 'If it is raining on my house two minutes from now, it will be raining on my house two minutes from now'. Although this tautological statement undermines the elegance of the logical deduction, it does nothing to increase the confidence ascribed to the overall hypothesis. Similarly, a hypothesis about attention and accuracy might be deduced, in part, from a strong-form 'near-tautological' proposition about attention and interest, but the confidence logically ascribed to the hypothesis remains a function of probabilities assigned to each of the underlying propositions. As

long as some subset of those probabilities are recognized by some subset of observers to be less than 100 percent (a condition that is virtually always satisfied with psychological propositions), results of a test of this hypothesis will be accordingly informative.¹

The Wallachs' Logical Analysis is Irrelevant to Most Hypotheses Concerning Psychological Processes

Thomas Edison once claimed that 'I have constructed three thousand different theories in connection with the electric light, each one of them reasonable and apparently likely to be true.' He added that he had disproved 2,998 of these theories by experiment (Grattan, 1933, p. 156). This quote illustrates a fundamental distinction that seems overlooked in the Wallachs' analysis: Logical deduction may tell us what *might be* going on in the real world, but it does not inform us as to what *really is* going on. Within psychology—as with any science—deduction is an indispensable tool for generating (or justifying) hypotheses, but it cannot persuade a scientist to transform conjecture into knowledge; epistemic alchemy demands data.

This is especially so when the objects of study are not merely psychological *phenomena* but psychological *processes*—processes that famously fail to operate strictly according to the prescriptions of logical philosophies. Deduction is helpful in deriving a prediction *that* a relation between variables will be observed, but is less helpful in explaining *why* that relation is observed. Nowadays, most psychological hypotheses not only predict relations, but predict also the psychological processes that account for those relations. As the Wallachs note, a predicted relation between social context and stereotype formation can be deduced from reasoning articulated by Hamilton and Gifford (1976), and also from reasoning articulated by Fiedler (1996). Although both lines of reasoning can predict exactly the same relation between contextual input and judgmental output, the two lines of reasoning comprise distinctly different hypotheses about underlying psychological processes. One or both hypotheses may be right, or both may be wrong. To figure it out demands some data. Data supporting one of these hypotheses will 'tell one nothing one could not have known already' (Wallach & Wallach, 1998, p. 191) only if one 'knew' already that the workings of the human mind matched exactly the deductive logic of the scientist. Few scientists assert such knowledge.

The Wallachs' Analysis May Have Counterproductive Consequences

We have shown that the Wallachs' analysis of psychological hypotheses is in error. But might they still be partially right? Even though they are clearly

falsifiable, might the logical structure of psychological hypotheses nonetheless contribute to the increasing perception that these hypotheses are 'safe yet boring' (Wegner, 1992)? Probably not. In our earlier article (Schaller et al., 1995) we argued that the interest value of an empirical study has little to do with the logical deducibility of the hypothesis it tests. In showing that virtually any hypothesis—including those with well-demonstrated interest value—can be deduced from (weak-form) 'near-tautologies', we showed how Latané and Darley's (1970) famous studies tested hypotheses that fit this putatively purposeless structure. The Wallachs amplify the point: They argue persuasively that the reason Latané and Darley's work generated such interest had nothing to do with its logical structure. Absolutely right.

The problem is not that tests of psychological hypotheses yield no knowledge about truth or falsity—they do—but rather that more and more hypotheses represent only small increments in conceptual conjecture. One reason for this trend lies in the evolving norms of the scientific culture. Higgins (1992) has argued that regulatory processes in the psychological sciences increasingly prescribe what scientists 'ought' to do, thus contributing to a mindset concerned with the avoidance of mistakes; correspondingly, there is an implicit discouragement of intellectual risk-taking. This shift has the predictable consequence of enhancing methodological complexity but debilitating conceptual creativity: 'to avoid the perception of mistakes, it is best to work within traditional boundaries—use conventional paradigms and interpret results in accordance with established theories' (Higgins, 1992, p. 491). In short, a prescriptive orientation defeats the spark of creativity, and so restricts the range of scientific ideas generated and tested.

Hull (1988) wrote that 'Without alternatives to be selected, scientific change cannot occur' (p. 254). Here we arrive at our fundamental objection to the Wallachs' analysis. By prescribing (on the basis of a logically unsound analysis) what psychologists ought and ought not be doing, their analysis may contribute to a mindset that stifles creative means of acquiring alternative visions. We worry that if readers accept their analysis uncritically, it may contribute unintentionally to the true cause of the dullness they detect.

If Read Critically, the Wallachs' Analysis Might Still Serve a Valuable Purpose

We have argued that the Wallachs' analysis is both wrong and wrongheaded; but it still may serve a valuable purpose if readers respond with thoughtful critical inquiry. Let us grant the Wallachs the role for which they have auditioned: Let us cast them as the voice of an impassioned minority, persistently expressing an unpopular belief. It may not matter that their belief fails to stand up to scrutiny. Like dance critics, persistent minorities

need not hold the right opinions in order to serve a positive function. Even if wrong and ultimately unpersuasive, persistent minorities do inspire others to question what they previously took for granted, and to do so critically and creatively (Nemeth, 1986). If that critical appraisal results in a rejection of the Wallachs' implicit prescriptions about what psychological scientists ought and ought not to do, then the Wallachs' articles might indeed contribute to our common goal of bolder, more creative science.

Note

1. A related consideration also reveals why implicit 'near-tautologies' fail to undermine the falsifiability of a hypothesis or the informativeness of research testing that hypothesis. In actual practice, any scientific experiment simultaneously tests a set of assumptions about conceptual relations and operationalizations (Duhem, 1914/1962). While there may be little empirical risk to the individual underlying assumptions, testing their veracity is rarely the purpose of the experiment. Instead, what is being tested is the entire scope of the theoretical argument—the specified manner in which the underlying assumptions fit together. Just as a set of individually lovely objects may be fit together to form an unlovely artwork, so too it is possible for individually accurate and/or 'obvious' psychological assertions to be assembled into a hypothesis that obviously fails to accurately describe actual psychological processes.

References

- Duhem, P. (1962). *The aim and structure of physical theory* (2nd ed.; P.P. Wiener, Trans.). New York: Atheneum. (Original work published 1914.)
- Fiedler, K. (1996). Explaining and simulating judgment biases as an aggregation phenomenon in probabilistic, multiple-cue environments. *Psychological Review*, *103*, 193–214.
- Grattan, C.H. (1933). Thomas Alva Edison: An American symbol. *Scribner's Magazine*, *94*, 151–156, 186–192.
- Hamilton, D.L., & Gifford, R.K. (1976). Illusory correlation in interpersonal perception: A cognitive basis of stereotypic judgments. *Journal of Experimental Social Psychology*, *12*, 392–407.
- Higgins, E.T. (1992). Increasingly complex but less interesting articles: Scientific progress or regulatory problem? *Personality and Social Psychology Bulletin*, *18*, 489–492.
- Hull, D.L. (1988). *Science as a process*. Chicago, IL: University of Chicago Press.
- Latané, B., & Darley, J.M. (1970). *The unresponsive bystander: Why doesn't he help?* New York: Appleton-Century-Crofts.
- Nemeth, C.J. (1986). Differential contributions of majority and minority influence. *Psychological Review*, *93*, 23–32.
- Schaller, M., Crandall, C.S., Stangor, C., & Neuberg, S.L. (1995). 'What kinds of social psychology experiments are of value to perform?' Comment on Wallach and Wallach (1994). *Journal of Personality and Social Psychology*, *69*, 611–618.

- Wallach, L., & Wallach, M.A. (1994). Gergen versus the mainstream: Are hypotheses in social psychology subject to empirical test? *Journal of Personality and Social Psychology*, 67, 233–242.
- Wallach, M.A., & Wallach, L. (1998). When experiments serve little purpose: Misguided research in mainstream psychology. *Theory & Psychology*, 8, 183–194.
- Wegner, D.M. (1992). The premature demise of the solo experiment. *Personality and Social Psychology Bulletin*, 18, 504–508.

MARK SCHALLER is an Assistant Professor of Psychology at the University of British Columbia. Previously, he held faculty appointments at the University of Texas at Arlington, and the University of Montana. In recent years his research has focused on statistical reasoning in everyday inference, processes of stereotype formation, and the psychological origins of cultural norms. ADDRESS: Department of Psychology, UBC, 2136 West Mall, Vancouver BC, Canada V6T 1Z4. [email: schaller@cortex.psych.ubc.ca]

CHRISTIAN S. CRANDALL, Associate Professor of Psychology at the University of Kansas, received his PhD from the University of Michigan in 1987. He is currently researching anti-fat attitudes, attribution and cross-cultural approaches to prejudice. He is also collecting data attempting convincingly to disconfirm the Duhem–Quine hypothesis. ADDRESS: Department of Psychology, 426 Fraser Hall, University of Kansas, Lawrence, Kansas, 66045, USA. [email: crandall@falcon.cc.ukans.edu]