

# FADS, FASHIONS, AND FOLDEROL IN PSYCHOLOGY<sup>1</sup>

MARVIN D. DUNNETTE

*University of Minnesota*

**T**HIS seemed to be a great idea when the 1965 Division 14 program was being arranged. When I chose the title, it seemed *very* clever, and I looked forward to venting my spleen a bit and to saying sage things about what is wrong with psychology and how its ills might be cured.

It was not long, however, before I felt misgivings about the whole enterprise and doubted whether I could meet the challenge to put up or to shut up that my own foolhardiness had cast upon me. At first I thought I knew the things in psychology that bothered *me*, but would there be consensus about this? Suddenly, I had the sobering thought that I might be the only one out of step and that, really, all was well with psychology.

So—seizing upon one of my own pet types of methodological folderol—I decided to run a survey! I contacted older or wiser heads,<sup>2</sup> to inquire what, if anything, was currently bothering them. Their responses might conservatively be described as a vast outpouring. The volume and intensity of their replies caused me to give up my speculations about joining the French Foreign Legion and show up here today after all.

With their suggestions, however, I was faced suddenly with a plethora of fads, fashions, and folderol and the need to make some systematic sense of them. Let me give you the flavor of my survey results by simply mentioning some of the things listed by me and by my respondents.

Fads—those practices and concepts characterized by capriciousness and intense, but short-lived interest—included such things as brainstorming, Q technique, level of aspiration, forced choice, critical incidents, semantic differential, role playing,

need theory, grids of various types, adjective checklists, two-factor theory, Theory X and Theory Y, social desirability, response sets and response styles, need hierarchies, and so on and so on.

Fashions—those manners or modes of action taking on the character of habits and enforced by social or scientific norms defining what constitutes the “thing to do”—included theorizing and theory building, criterion fixation, model building, null hypothesis testing, sensitivity training, being productive at work, developing authentic relationships, devising “cute” experiments, simulation, using “elegant” statistics, and so on.

Finally, folderol—those practices characterized by excessive ornamentation, nonsensical and unnecessary actions, trifles and essentially useless and wasteful fiddle-faddle—included tendencies to be fixated on theories, methods, and points of view, conducting “little” studies with great precision, attaching dramatic but unnecessary trappings to experiments, asking unimportant or irrelevant questions, grantsmanship, coining new names for old concepts, fixation on methods and apparatus, seeking to “prove” rather than “test” theories, and myriad other methodological ceremonies conducted in the name of rigorous research.

But, even armed with my list, about all I could say is that there are many things going on in psychology that reasonably responsible people were willing to label faddish folderol. It accomplished the aim of identifying some of the less honorable things we all are doing, but it seemed rather sterile as a source of prescriptive implications. What was needed was a better taxonomy for listing psychology’s ills than the rather artificial trichotomy established by my title.

One approach might be through some form of cluster analysis—but my data did not prove amenable to any of the widely used methods such as Pattern Analysis, Elementary Linkage Analysis, Hierarchical Linkage Analysis, Hierarchical Syndrome Analysis, Typal Analysis, Rank Order Typal Analysis, Comprehensive Hierarchical Analysis, or even Multiple Hierarchical Classification.

<sup>1</sup> Invited Address presented to Division 14 at American Psychological Association, Chicago, September 1965.

<sup>2</sup> I should like to thank the following persons for their readiness to come to my aid, but please understand that I do so only because of a strong sense of gratitude and with no thought of having them join me out at the end of the limb. They are: David P. Campbell, John P. Campbell, Alphonse Chapanis, Edwin E. Ghiselli, Mason Haire, James Jenkins, Quinn McNemar, Paul Meehl, William A. Owens, Jr., Bernard Rimland, Auke Tellegen, Rains Wallace, and Karl Weick.

Just when I was facing this impasse, I received the best-selling book by Eric Berne (1964) titled *Games People Play*. I was stimulated by this magnificent book to give thought to the games psychologists play. Somewhat to my surprise, I experienced no difficulty slipping into the robes of a medical clinician—intent on describing fully and completely the behavioral symptomatology of psychology's distress in terms of the games we all play—many of which may reflect underlying pathologies leading us down the primrose path to non-science.

The games can be discussed under six broad headings—The Pets We Keep; The Fun We Have; The Names We Love; The Delusions We Suffer; The Secrets We Keep; and The Questions We Ask.

#### THE PETS WE KEEP

Subtitled "What Was Good Enough for Daddy Is Good Enough for Me," this game is characterized by an early and premature commitment to some Great Theory or Great Method. One major effect is to distort research problems so that they fit the theory or the method. The theory, method, or both can be viewed as pets inherited by fledgling psychologists and kept and nurtured by them, in loving kindness, protecting them from all possible harm due to the slings and arrows and attacks from other psychologists who, in turn, are keeping their own menageries.

At a general level, the premature commitment to a theory is usually accompanied by the set to *prove* rather than to modify the theory. The problem and its potentially bad outcome was outlined years ago by T. C. Chamberlin (1965), a well known geologist. He stated:

The moment one has offered an original explanation for a phenomenon which seems satisfactory, that moment affection for his intellectual child springs into existence; . . . there is an unconscious selection and magnifying of the phenomena that fall into harmony with the theory and support it, and an unconscious neglect of those that fail of coincidence. . . . When these biasing tendencies set in, the mind rapidly degenerates into the partiality of paternalism. . . . From an unduly favored child, it [the theory] readily becomes master, and leads its author whithersoever it will [p. 755].

It is not difficult in psychology to recognize the sequence of events described by Chamberlin. A pessimist might, in fact, find it difficult to identify any psychological theories which do *not* currently enjoy this form of affectionate nurturing. On the

other hand, a more optimistic view might accord to theories the important function of ordering and systematizing the conduct of research studies. What is to be avoided, of course, is the kind of paternal affection and closed mind described by Chamberlin.

The problem in psychology is made more severe, however, by the inexplicitness (Fiegel, 1962) and, as Ritchie<sup>3</sup> has called it, the "incurable vagueness" with which most theories are stated—but then, it should be clear that vagueness in theory construction may simply be part of the game, insuring higher likelihood of a pet theory's long life.

Methodologically, our favored *pets* include factor analysis, complex analysis of variance designs, the concept of statistical significance, and multiple-regression analyses. It is common for psychologists to apply so-called sophisticated methods of analysis to data hardly warranting such careful attention. I shall not try to enumerate the nature of the painstaking activities included in the game of statistical pet keeping. I refer those of you who are interested to excellent papers by McNemar (1951) and by Guilford (1960). The net effect, however, is that attention to relevant and important scientific questions is diminished in favor of working through the subtle nuances of methodological manipulation. As my colleague, David Campbell remarked,<sup>4</sup>

We seem to believe that TRUTH will be discovered somehow through using more and more esoteric techniques of data manipulation rather than by looking for it in the real world.

Or, as Platt (1964) has said:

Beware of the man of one method or one instrument, either experimental or theoretical. He tends to become method-oriented rather than problem oriented; the method oriented man is shackled [p. 351].

#### THE NAMES WE LOVE

An alternate title for this game is "What's New Under the Sun?" Unfortunately, an undue amount of energy is devoted to the Great Word Game—the coining of new words and labels either to fit old concepts or to cast new facts outside the ken of a theory in need of protection.

Just one from among many possible examples is the great emphasis in recent years on Social Desir-

<sup>3</sup> B. F. Ritchie, unpublished work.

<sup>4</sup> Personal communication, 1965.

ability—a new label for a phenomenon in test-taking behavior dealt with extensively by Meehl and Hathaway (1946), Jurgensen, and others many years previously, but which did not create much interest because they failed at the time to coin a label sufficiently attractive to “grab” other psychologists.

As Maier (1960) has so aptly pointed out, one major effect of the Name Game is to sustain theories even if the facts seem to refute them. If facts appear that cannot be ignored, relabeling them or renaming them gives them their own special compartment so that they cease to infringe upon the privacy of the theory.

Perhaps the most serious effect of this game is the tendency to apply new names in psychological research widely and uncritically before sufficient work has been done to specify the degree of generality or specificity of the “trait” being dealt with. Examples of this are numerous—*anxiety*, *test-taking anxiety*, *rigidity*, *social desirability*, *creativity*, *acquiescence*, *social intelligence*, and so on—*ad infinitum*.

#### THE FUN WE HAVE

A suitable title for this game would be—quite simply—“Tennis Anyone?” But the game has many variants, including *My Model Is Nicer than Your Model!*, *Computers I Have Slept With!*, or the best game of all—*A Difference Doesn't Need to Make a Difference if It's a Real Difference*.

As should be clear, the underlying theme of the game—Tennis Anyone?—is the compulsion to forget the problem—in essence to forget what we are really doing—because of the fun we may be enjoying with our apparatus, our computers, our models or the simple act of testing statistical null hypotheses. Often, in our zest for this particular game, we forget not only the problem, but we may even literally forget to look at the data!

The most serious yet most common symptom of this game is the “glow” that so many of us get from saying that a result is “statistically significant.” The song and dance of null hypothesis testing goes on and on—apparently endlessly. In my opinion, this one practice is as much responsible as anything for what Sommer (1959) has called the “little studies” and the “little papers” of psychology.

As so many others have pointed out (Binder, 1963; Grant, 1962; Hays, 1964; Nunnally, 1960;

Rozeboom, 1960), the major difficulty with psychology's use of the statistical null hypothesis is that the structure of scientific conclusions derived thereby is based on a foundation of triviality. When even moderately small numbers of subjects are used nearly all comparisons between means will yield so-called “significant” differences. I believe most psychologists will agree, in their more sober and less fun-loving moments, that small differences and inconsequential correlations do not provide a sufficient yield either for accurately predicting other persons' behavior or for understanding theoretically the functional relations between behavior and other variables. Yet, most of us still remain content to build our theoretical castles on the quicksand of merely rejecting the null hypothesis.

It may seem that my criticism of this particular game is unduly severe. Perhaps the differences reported in our journals are not really all that small. In order to examine this question, I asked one of my research assistants, Milton Hakel, to sample recent issues of four APA Journals—the *Journal of Applied Psychology*, *Journal of Abnormal and Social Psychology*, *Journal of Personality and Social Psychology*, and *Journal of Experimental Psychology*. He selected randomly from among studies employing either *t* tests or complex analysis of variance designs, and converted the *t* or *F* values to correlation ratios ( $\eta$ ) in order to estimate the strength of association between independent and dependent variables.

The distribution of the 112 correlation ratios ranged from .05 to .92 with a median value of .42. Five percent of the studies showed values below .20; over one-sixth were below .25; and nearly one-third failed to reach .30. The only encouragement I derive from these data stems from my identification with industrial psychology. At a time when many in industrial psychology are worried because predictive validities rarely exceed .50, it is at least reassuring—though still disconcerting—to note that our brethren in social and experimental psychology are doing little better.

It is particularly informative to note the conclusions made by the authors of the articles sampled by Hakel. Authors of the study yielding the  $\eta$  of .05 concluded “that rating-scale format is a determiner of the judgment of raters in this sample [Madden & Bourden, 1964].” In an investigation yielding an  $\eta$  of .14, the authors con-

cluded "that highly creative subjects give the greatest number of associations and maintain a relatively higher speed of association throughout a 2 minute period [Mednick, Mednick, & Jung, 1964]."

Surprisingly, these rather definite conclusions differ little in tone from those based on studies yielding much stronger relationships. For example, a study yielding an eta of .77 is summarized with "Highly anxious subjects tended to give sets of word associates higher in intersubject variability than nonanxious subjects [Brody, 1964]." In like manner, the conclusion stated for a study yielding an eta of .63 was simply "It was found that reinforcement affected subjects' verbalizations [Ganzer & Sarason, 1964]."

It seems abundantly clear that our little survey provides convincing and frightening evidence that playing the game of null hypothesis testing has led a sizeable number of psychologists to lose sight of the importance of the strength of relationships underlying their conclusions. I could not agree more fully with Nunnally (1960), who has said:

it would be a pity to see it (psychology) settle for meager efforts . . . encouraged by the use of the hypothesis testing models. . . . We should not feel proud when we see the psychologist smile and say "the correlation is significant beyond the .01 level." Perhaps that is the most that he can say, but he has no reason to smile [p. 650].

#### THE DELUSIONS WE SUFFER

This is probably the most dangerous game of all. At the core, it consists of maintaining delusional systems to support our claims that the things we are doing *really* constitute good science. The game develops out of a pattern of self-deceit which becomes more ingrained and less tractable with each new delusion. Thus, an appropriate subtitle is "This Above All, to Thine Ownself Be *False!*"

The forms of these delusions are so numerous and so widespread in psychology that time permits only brief mention of a few.

One common variant of the game can be called, "Boy, Did I Ever Make Them Sit Up and Take Notice!" The argument is often made and seemingly almost always accepted that if a new theory or method stimulates others to do research, it *must* be good. Although I greatly dislike analogic arguments, I am compelled to suggest that such reasoning is very similar to stating that accidental

fire must be good simply because it keeps so many firemen busy. Unfortunately, an inestimable amount of psychological research energy has been dissipated in fighting brush fires spawned by faddish theories—which careful research might better have refuted at their inception.

It is probably far too much to hope that we have seen the last of the studies "stimulated" by Sheldon's notions about physique and temperament, or by the overly simplified but widely popular two-factor theory of job motivation (Herzberg, Mausner, & Snyderman, 1959).

A second common delusion seems to arise out of the early recognition that gathering data from real people emitting real behaviors in the day-to-day world proves often to be difficult, unwieldy, and just plain unrewarding. Thus many retreat into the relative security of experimental or psychometric laboratories where new laboratory or test behaviors may be concocted to be observed, measured, and subjected to an endless array of internal analyses. These usually lead to elaborate theories or behavioral taxonomies, entirely consistent within themselves but lacking the acid test of contact with reality. Last year, McNemar (1964) summarized once more for us the evidence showing the pathetic record of factor analytically derived tests for predicting day-to-day behavior. A former professor at Minnesota used to say—when describing a lost soul—"He disappeared into the Jungle of Factor Analysis—never to be heard from again." Psychologists who choose to partake of the advantages of the more rigorous controls possible in the psychometric or experimental laboratories must also accept responsibility for assuring the day-to-day behavioral relevance of the behavioral observations they undertake.

A third unfortunate delusion rationalizes certain practices on the grounds that they are intrinsically good for humanity and that they need not, therefore, meet the usual standards demanded by scientific verification. In this regard, Astin (1961) has done an effective job of analyzing the functional autonomy of psychotherapy and offers a number of reasons why it continues to survive in spite of a lack of evidence about its effectiveness. In industrial psychology, a most widespread current fashion is the extensive use by firms of group-process or sensitivity-training programs; the effectiveness of such programs is still proclaimed solely on the basis of testimonials, and a primary ra-

tionale for their inadequate evaluation is that they are a form of therapy and must, therefore, be good and worthwhile.

Finally, yet another pair of delusions, representing polar opposites of one another, were discussed by Cronbach (1957) in his American Psychological Association Presidential Address. One extreme, shown chiefly by the experimentalists, treats individual differences as merely bothersome variation—to be reduced by adequate controls or treated as error variance in the search for General Laws. Such assumptions cannot help but lead to an oversimplified image of man, for the simplification is introduced at the very beginning. We cannot expect a science of human behavior to advance far until the moderating effects of individual variation on the functional relationships being studied are taken fully into account. People do, after all, differ greatly from one another and they differ even more from monkeys, white rats, or pigeons. It should not *really* be too heretical to suggest that many of the lawful relations governing the behavior of lower organisms may be inapplicable to the human species and, moreover, that laws describing the behavior of certain selected human subjects—such as psychology sophomores—may upon examination prove only weakly applicable to many other individuals. It should be incumbent upon the experimentalist or the theorist either to incorporate a consideration of individual differences into his research and theorizing or to define explicitly the individual parameters or population characteristics within which he expects his laws to be applicable.

The other extreme, actually extending considerably beyond the correlational psychology discussed by Cronbach, is just as delusory and even more detrimental to the eventual development of psychology than the one just discussed. Differences between individuals are regarded as so pervasive that it is assumed *no* laws can be stated. The likely outcome of a strong commitment to this point of view must ultimately be an admission that the methods of science cannot be applied to the study of human behavior. Yet, this outcome is not often openly recognized or honestly accepted by those believing in the ultimate uniqueness of each individual. Instead, they speak of “new approaches,” less “mechanistic emphases,” and a more “humanistic endeavor.”

Cronbach, nearly a decade ago, sounded an

urgent call for his fellow psychologists to cast aside the delusions represented by these two extremes. Unfortunately, today we seem no closer to achieving this end than we were then.

#### THE SECRETS WE KEEP

We might better label this game “Dear God, Please Don’t Tell Anyone.” As the name implies, it incorporates all the things we do to accomplish the aim of looking better in public than we really are.

The most common variant is, of course, the tendency to bury negative results. I only recently became aware of the massive size of this great graveyard for dead studies when a colleague expressed gratification that only a third of his studies “turned out”—as he put it.

Recently, a second variant of this secrecy game was discovered, quite inadvertently, by Wolins (1962) when he wrote to 37 authors to ask for the raw data on which they had based recent journal articles. Wolins found that of 32 who replied, 21 reported their data to be either misplaced, lost, or inadvertently destroyed. Finally, after some negotiation, Wolins was able to complete seven re-analyses on the data supplied from 5 authors. Of the seven, he found gross errors in three—errors so great as to clearly change the outcome of the results already reported. Thus, if we are to accept these results from Wolins’ sampling, we might expect that as many as one-third of the studies in our journals contain gross miscalculations. In fact, this variant of the secrecy game might well be labeled “I Wonder Where the Yellow (data) Went.” In commenting on Wolins’ finding, Friedlander (1964), impressed by the strong commitments psychologists hold for their theories, tests and methods, suggests that “Hope springs eternal—and is evidently expressed through subjective arithmetic”—a possibility which is probably too close to the truth to be taken lightly.

Another extremely vexing and entirely unnecessary type of secrecy is clearly apparent to anyone who takes but a moment to page through one of our current data-oriented psychological journals. I chose a recent issue of the *Journal of Personality and Social Psychology*. It was very difficult to find such mundane statistics as means or standard deviations. Instead, the pages abounded with analysis of variance tables, charts, *F* ratios, and even *t* tests in the absence of their corresponding means

and *SDs*. The net effect of this is to make very difficult and often impossible any further analyses that a reader might want to undertake. The implication of this, it seems to me, is that many authors have actually failed to bother computing such statistics as means or *SDs* and that, further, they probably have not examined their data with sufficient care to appreciate in any degree what they may really portray.

Other examples of the secrecy game abound. They include such practices as dropping subjects from the analyses—a practice discussed at some length in the critical review of a sampling of dissonance studies by Chapanis and Chapanis (1964), experimenter-biasing factors, incomplete descriptions of methodology, failure to carry out or to report cross-validation studies, and the more general problem of failure to carry out or to report replication studies.

I believe you will agree that these tactics of secrecy can be nothing but severely damaging to any hopes of advancing psychology as a science. It seems likely that such practices are rather widely applied in psychology by psychologists. I suggest that we vow here and now to keep these secrecy games secret from our colleagues in the other sciences!

#### THE QUESTIONS WE ASK

There are many titles that might be appropriate for this last game that I shall discuss. One might be, "Who's on First?"—or better yet, "What Game Are We In?"—or a rather common version in these days of large Federal support for research, "While You're Up, Get Me a Grant." My major point here is quite simply that the other games we play, the pets we keep, our delusions, our secrets, and the Great Name Game interact to cause us to lose sight of the essence of the problems that need to be solved and the questions that need answers. The questions that get asked are dictated—all too often—by investigators' pet theories or methods, or by the need to gain "visibility" among one's colleagues. One of my respondents—a younger but undoubtedly wiser head than I—summed it up nicely.<sup>5</sup> He said:

Psychologists seem to be afraid to ask really important questions. The whole Zeitgeist seems to encourage research efforts that earn big grants, crank out publications frequently and regularly, self-perpetuate themselves, don't

entail much difficulty in getting subjects, don't require the researchers to move from behind their desks or out of their laboratories except to accept speaking engagements, and serve to protect the scientist from all the forces that can knock him out of the secure "visible circle."

Another of my respondents, a verbal behavior researcher, illustrated the dilemma by mentioning a fellow researcher who phrased his research question as: "How do the principles of classical and instrumental conditioning explain the learning of language?" This sort of question is clearly illustrative of the tendency to defer too readily to existing popular points of view and to allow them to distort the direction of research activities. It would be better simply to ask "What is learned?" rather than making the premature assumptions that (a) language is learned in the sense that the term *learning* is usually used or (b) *all* learning is of only two types.

An even more serious and, unfortunately, probably more common form of the question-asking game is the game of "Ha! Sure Slipped That One Past You, Didn't I?" Here, the investigator shrewdly fails to state the question he is trying to answer, gathers data to provide answers to simpler questions, and then behaves as if his research has been relevant to other unstated but more important and more interesting problems. The vast majority of studies devoted to measuring employee attitudes have committed this error. It is no trick to develop questionnaires to gather systematically the opinions of workers about their jobs. It is quite something else, however, suddenly to begin talking about measures of employee *motivation* and to suggest that the employee responses have direct relevance to what they may actually do on the job. The literature on response set and response style is another clear case of new questions being designed to fit existing answers. Showing high correlations between scores on empirically developed scoring keys and the numbers of items keyed *True* or some other item index should not be taken as having any bearing on the empirical validities of these keys. Yet for over a decade, our literature has been burdened with all sorts of set and style studies characterized by seemingly endless factor analyses and silly arguments between persons committed to acquiescence and those committed to social desirability.

Thus, we all are far too eager to ask such questions as: "What problems can be *easily* answered?"

<sup>5</sup> J. P. Campbell, personal communication, 1965.

"What else can I do with my test?" "What problems or questions does my theory lead to?" "What aspects of behavior can I study with my computer or with my apparatus?" or "What problems can I find that I can fit this method to?"

Certainly, as psychologists—as scientists presumably interested in the subject matter of human behavior—we should be able to do better than this!

#### THE CAUSES

You may have inferred by now that I feel some sense of pessimism about the current state of psychology. Based on what you have heard so far, such an inference is probably appropriate. I *do* believe that the games I have described offer little that can be beneficial for psychology in the long run. The behaviors underlying the games represent enormous and essentially wasteful expenditures of our own research energy.

Even so, my mood is not basically pessimistic. In fact, we should be able to emerge from this soul searching with a constructive sense of discontent rather than one of destructive despair. The description of our condition carries with it a number of implications for corrective action. Moreover, we may infer from the condition some possible causes, and listing them should also suggest possible correctives.

To this end, let me consider briefly what I believe to be the major causes of psychology's fads, fashions, and folderol.

The most important, I believe, is related quite directly to the relative insecurity of being a scientist, a problem that is particularly acute in psychology where we must cope with such complex phenomena as those involved in the study of human behavior. The scientist's stance includes the constant need to doubt his own work. Moreover, the long-range significance of his work cannot often be forecast, and rarely can the scientist—least of all, perhaps, the psychologist—preplan his inspirations and his ideas. It is little wonder, then, that many seek, through their theories, methodologies, or other of the games we have discussed, to organize, systematize, and regularize their creative output. When viewed against the backdrop of publication pressures prevailing in academia, the lure of large-scale support from Federal agencies, and the presumed necessity to become "visible" among one's colleagues, the insecurities of undertaking

research on important questions in possibly untapped and unfamiliar areas become even more apparent. Stern (1964) has recently very effectively stated the case for the desirability of moving from research into equally fulfilling careers of teaching and administration. But we cannot forget that the value system of science places research and publication at the peak, and it should, therefore, be no surprise that the less able researchers in psychology—learning early that no great breakthrough is in the offing—simply seek to eat their cake and have it too, by playing the games and the song and dance of scientific research, usually convincing even themselves that the games are "for real" and that their activities really "make a difference."

The perpetuation of this state of affairs is related to our present system of graduate education. Many psychology graduate students today find themselves under the tutelage of a faculty member who has bought the system wholeheartedly. Such students live for a period of from 3 to 8 years in an environment that enforces and reinforces the learning of a particular approach, a narrow point of view or a set of pet methodologies which come to define for them the things they will pursue as psychologists.

#### THE REMEDY

But here I am—sounding pessimistic and noxious again, and getting farther out on the limb than I really want to be.

In order to convince you of my good intentions and my hope for the future, I had better get on with some constructive suggestions. My suggested remedy—if it can be called that, for indeed it may be more painful than the disease—can be summarized in five imperative statements:

1. Give up constraining commitments to theories, methods, and apparatus!
2. Adopt methods of multiple working hypotheses!
3. Put more eclecticism into graduate education!
4. Press for new values and less *pretense* in the academic environments of our universities!
5. Get to the editors of our psychological journals!

Let me elaborate briefly on each of these recommendations.

First, I advocate a more careful and studied choice of research questions. As should be apparent, I believe research energy should be directed toward questions that contain as few as possible of any prior unproven assumptions about the nature of man. We must be constantly alert to the narrowing of research perspectives due to prior theoretical or methodological commitments. I am calling for less premature theorizing—particularly that which leads to vaguely stated “wide-band” theories that are often essentially incapable of disproof.

I am not advocating the abandonment of deduction in psychology; in fact, psychology needs stronger and more specific deductions rather than the weak and fuzzy ones so typical of so many current theories. What I am advocating is the more systematic study of lawful relationships *before* interpretations are attempted. When explanation *is* attempted, the data should be sufficient to allow hypotheses to be stated with the clarity and precision to render them directly capable of disproof. As the philosopher Karl Popper has said, there is no such thing as proof in science; science advances only by disproofs.

This leads directly to my second recommendation which is to state and systematically test multiple hypotheses. Platt recently advocated this approach which he calls *Strong Inference* (Platt, 1965). The approach entails devising multiple hypotheses to explain observed phenomena, devising crucial experiments each of which may exclude or disprove one or more of the hypotheses, and continuing with the retained hypotheses to develop further subtests or sequential hypotheses to refine the possibilities that remain. This process does not seem new; in fact it is not. It simply entails developing ideas or leads, stating alternative possibilities, testing their plausibility, and proceeding to develop predictive and explanatory evidence concerning the phenomena under investigation. One might say that the research emphasis is one of “studying hypotheses” as opposed to “substantiating theories.” The difference seems slight, but it is really quite important. However, in psychology, the approach is little used, for, as we have said, the commitments are more often to a theory than to the process of *finding out*.

The method of multiple hypotheses takes on greatly added power when combined with greater care in the analysis and reporting of research re-

sults. Instead of serving as the sole statistical test of hypotheses, the statistical null hypothesis should always be supplemented by estimates of strength of association. The psychologist owes it to himself to determine not only whether an association exists between two variables—an association which may often be so small as to be trivial—but also to determine the probable magnitude of the association. As Hays (1964) has suggested, if psychologists are content to adopt conventions (such as .05 or .01) for deciding on statistical significance, they should also adopt conventions concerning the strength of association which may be sufficiently large to regard as worthy of further investigation. Obviously, such conventions cannot be the same for all areas and for all research questions, but it should be clear that an emphasis on magnitude estimation will demand that researchers give much more careful thought than they now do to defining ahead of time the actual magnitudes that will be regarded as possessing either theoretical or practical consequence.

By now, it is apparent why my fifth recommendation has to do with our journals. It will require a new kind of surveillance from both the editors and their consultants if we are to implement the greater care in research conception and in data analysis and reporting that I am advocating. When and if null-hypothesis testing is accorded a lower position in the status hierarchy and comes to be supplemented by emphases on Strong Inference and magnitude estimation, I would predict that the bulk of published material will, for a time, greatly diminish. That which does appear, however, will be guaranteed to be of considerably greater consequence for furthering our understanding of behavior.

One of the possible loopholes in the method of Strong Inference, it should be clear, is the great difficulty of designing and carrying out crucial experiments. Recently Hafner and Presswood (1965) described how faulty experiments had led physicists astray for several decades as they sought to explain the phenomenon of beta decay. We must broaden our conception of multiple hypotheses to include as one quite plausible hypothesis the possibility of poorly conceived or poorly conducted experiments. This, of course, simply speaks to the need for more replication in psychology of crucial experiments, a practice which undoubtedly would become more widespread if



psychologists possessed fewer of their own theoretical pets and stronger motivation to examine systematically whole sets of contending hypotheses and alternative explanations.

My third and fourth recommendations need not be elaborated extensively. Both are intended to foster less pretense in the conduct of psychological research by enabling those scholars who may be ill fitted for the research enterprise to gain rewards in other endeavors. The change in the academic atmosphere would need to take the form of according more status to good teaching and to good administration. Perhaps this change would be most rapidly fostered if the scientific games I have described would be more readily recognized for what they are and appropriately devalued in the scheme of things within academia.

Obviously, greater eclecticism in graduate education is crucial to the successful outcome of my other suggestions. It is difficult to know how this can be implemented. But, at least, the goals seem clear. We desire to teach the core of psychology's knowledge and methods, its subject matter and its questions, the statistical methods and their appropriate applications—but most of all, through selection or training or both, we should seek to turn out persons with intense curiosity about the vast array of psychological questions and problems occurring everywhere in the world around us, with a willingness to ask *open* questions unhampered by the prior constraints of a particular point of view or method. Let us hope that graduate education, in the years ahead, will become more eclectic and that even the Great Men in our field may adopt a sense of humility when transmitting knowledge to the fledglings of our science.

#### THE OUTCOME: UTOPIA

How do I envision the eventual outcome if all these recommendations were to come to pass? What would the psychologizing of the future look like and what would psychologists be up to?

Chief among the outcomes, I expect, would be a marked lessening of tensions and disputes among the Great Men of our field. I would hope that we might once again witness the emergence of an honest community of scholars all engaged in the zestful enterprise of trying to describe, understand, predict, and control human behavior.

Certainly our journals would be more meaty and less burdensome. There would be more honesty

in publishing the fruits of one's labors. Negative results—the disproof of theoretical formulations and the casting aside of working hypotheses—would be a more important part of the journals' contents. In consequence, the journals would contribute more meaningfully to the broad effort to achieve understanding, and we should expect to witness a sharp decline in the number of disconnected little studies bearing little or no relation to each other.

Moreover, I expect that many present schisms in psychology would be welded. The academic-professional bipolarity described by Tryon (1963) would be lessened, for the advantages to both of close association between basic researchers and those practicing the art of psychology should become more apparent. The researchers would thereby establish and maintain contact with the real world and real problems of human behavior, and the professional practitioners would be more fully alert to the need for assessing their methods by generating and testing alternate deductions and hypotheses growing out of them.

Thus, in the long run we might hope for fewer disputes, a spirit of more open cooperation, greater innovation in the generation and testing of working hypotheses, greater care and precision in the development of theoretical formulations, and increased rigor in specifying the magnitude of outcomes such that they have both practical and theoretical importance.

Does this sound like Utopia? Indeed it does. But is it too much to expect of a science now well into its second 100 years? I think not. Let us get on then with the process of change and of reconsolidation.

#### REFERENCES

- ASTIN, A. W. The functional autonomy of psychotherapy. *American Psychologist*, 1961, 16, 75-78.
- BERNE, E. *Games people play*. New York: Grove Press, 1964.
- BINDER, A. Further considerations on testing the null hypothesis and the strategy and tactics of investigating theoretical models. *Psychological Review*, 1963, 70, 107-115.
- BRODY, N. Anxiety and the variability of word associates. *Journal of Abnormal and Social Psychology*, 1964, 68, 331-334.
- CHAMBERLIN, T. C. The method of multiple working hypotheses. *Science*, 1965, 148, 754-759.
- CHAPANIS, N., & CHAPANIS, A. Cognitive dissonance: Five years later. *Psychological Bulletin*, 1964, 61, 1-22.

- CRONBACH, L. J. The two disciplines of scientific psychology. *American Psychologist*, 1957, **12**, 671-684.
- FEIGEL, H. Philosophical embarrassments of psychology. *Psychologische Beiträge*, 1962, **6**, 340-364.
- FRIEDLANDER, F. Type I and Type II bias. *American Psychologist*, 1964, **19**, 198-199.
- GANZER, V. J., & SARASON, I. G. Interrelationships among hostility, experimental conditions, and verbal behavior. *Journal of Abnormal and Social Psychology*, 1964, **68**, 79-84.
- GRANT, D. A. Testing the null hypothesis and the strategy and tactics of investigating theoretical models. *Psychological Review*, 1962, **69**, 54-61.
- GUILFORD, J. P. Psychological measurement a hundred and twenty-five years later. Invited address presented at American Psychological Association, Chicago, September 1960.
- HAFNER, E. M., & PRESSWOOD, S. Strong inference and weak interactions. *Science*, 1965, **149**, 503-510.
- HAYS, W. L. *Statistics for psychologists*. New York: Holt, Rinehart & Winston, 1964.
- HERZBERG, F., MAUSNER, B., & SNYDERMAN, B. *The motivation to work*. New York: Wiley, 1959.
- MADDEN, J. M., & BOURDON, R. D. Effects of variations in rating scale format on judgment. *Journal of Applied Psychology*, 1964, **48**, 147-151.
- MAIER, N. R. F. Maier's law. *American Psychologist*, 1960, **15**, 208-212.
- MCNEMAR, Q. Lost our intelligence? Why? *American Psychologist*, 1964, **19**, 871-882.
- MEDNICK, M., MEDNICK, S. A., & JUNG, C. C. Continual association as a function of level of creativity and type of verbal stimulus. *Journal of Abnormal and Social Psychology*, 1964, **69**, 511-515.
- MEHL, P. E., & HATHAWAY, S. R. The K factor as a suppressor variable in the MMPI. *Journal of Applied Psychology*, 1946, **30**, 525-564.
- NUNNALLY, J. The place of statistics in psychology. *Educational and Psychological Measurement*, 1960, **20**, 641-650.
- PLATT, J. R. Strong inference. *Science*, 1964, **146**, 347-352.
- ROZEBOOM, W. W. The fallacy of the null-hypothesis significance test. *Psychological Bulletin*, 1960, **57**, 416-428.
- SOMMER, R. On writing little papers. *American Psychologist*, 1959, **14**, 235-237.
- STERN, C. Thoughts on research. *Science*, 1964, **148**, 772-773.
- TRYON, R. C. Psychology in flux: The academic-professional bipolarity. *American Psychologist*, 1963, **18**, 134-143.
- WOLINS, L. Responsibility for raw data. *American Psychologist*, 1962, **17**, 657-658.