REVIEW of PERSONALITY and SOCIAL PSYCHOLOGY

Edited by
LADD WHEELER

1

Richard D. Ashmore
Brian Caddick
Francis K. Del Boca
Bella M. De Paulo
Ronald C. Dillehay
Kenneth J. Gergen
G. P. Ginsburg
Jeanne Longley
Colin Martindale
David C. McClelland
Jill Morawski
Michael T. Nietzel
Dean G. Pruitt
Robert Rosenthal
Carin Rubenstein
Phillip Shaver
Gudmund J.W. Smith
Bert Westerlundh
Jerry S. Wiggins
Miron Zuckerman

Published in cooperation with the Society for Personality and Social Psychology
(Division 8 of the American Psychological Association)

REVIEW of PERSONALITY and SOCIAL PSYCHOLOGY: 1

REVIEW OF PERSONALITY AND SOCIAL PSYCHOLOGY

REVIEW of PERSONALITY and SOCIAL PSYCHOLOGY

1

Edited by
LADD WHEELER

Published in cooperation with the SOCIETY FOR PERSONALITY AND SOCIAL PSYCHOLOGY (Division 8, American Psychological Association)

For information address:

SAGE PUBLICATIONS, INC.
275 South Beverly Drive
Beverly Hills, California 90212

SAGE PUBLICATIONS LTD
28 Banner Street
London EC1Y 8QE, England

Printed in the United States of America

International Standard Book Number 0-8039-1457-1 (hardcover)
0-8039-1458-X (softcover)

International Standard Serial Number 0270-1987

FIRST PRINTING

CONTENTS

Editor's Introduction

This new annual series was originated by the Society for Personality and Social Psychology (Division 8 of the American Psychological Association) as a special means of international communication. A current problem in our field is that personality and social psychologists do not know one another's work. It is true that they read some of the same journals, but each group also has its more specialized journals and book series, and the information overload can be impossible. As a social psychologist, I am not likely to find the time to read many long and detailed empirical studies in personality. But I *do* want to know what is happening in that closely allied field, and I want my students to know. For example, in this first volume, I learned more about personality measurement from McClelland's chapter than I would have believed possible from such a small investment of time. And my view of personality theory was drastically changed by Wiggins's chapter. I believe a similar benefit will be gained by readers who are primarily steeped in the literature of personality—and who are likely to find valuable insights from social psychologists in the pages that follow.

A second problem is that we North Americans aren't very much aware of advances made in other places. Only a handful of American psychologists are aware of Gudman Smith's 20-year research program on perceptgenesis (described in this volume). By having a distinguished international Editorial Board, we hope to help reduce this type of provincialism.

In a letter to me after I was named Editor, Irwin Altman wrote: "I feel that the *Review* should not only be a summary and integration of established areas of personality and social psychology, but I would very much like to see it receptive to emerging areas that have a sufficient base of empirical and theoretical work. In this respect, the *Review* can serve as *the* main evolutionary vehicle of the field . . . ." I agreed with Irv then, I agree now—let us try to achieve these goals. In our endeavor, we welcome the suggestions of our readers and colleagues in the areas of personality and social psychology for topics and authors (both in North America and abroad).

The *Review* will not be limited to any particular type of article, and—as you will see—there are many types in this first volume. We will always try, however, to focus on the frontiers of theory and method. We have tried to select articles of fascination to any serious graduate student or professional, and the articles have been written without the assumption of extensive background in the area. The decision of the Society to publish the *Review* in both soft- and hardcover editions is consistent with our belief that it should be read, marked, and carried around by a lot of people (not just the well-heeled).

No idea is the product of a single individual—and the concepts which led to the *Review's* debut grew out of the thoughts of a number of individuals, including Harry Triandis, Jerry Clore, James H. Davis, Ross Parke, Bibb Latané, Seymour Rosenberg, Irwin Altman, and Clyde Hendrick. I have profited from their suggestions. Another set of intellectual debts is certainly owed to the Fellows of the Society (who were generous with both their suggestions for articles, and for possible Editorial Board members); to the Editorial Board; and to the following individuals who assisted in the article review process: John Brigham, Robert C. Carson, Kenneth Craik, Germaine de Montmollin, Robert S. Feldman, G. P. Ginsburg, Irving Janis, Eric Klinger, Walter Mischel, Reidar Ommundson, Harry Reis, William McKinley Runyan, and Edward Sampson. The Associate Editors were more than helpful . . . they were wonderful! And I (as well as the Publishers) am indebted to Barbara Fox for assistance with the cover design.

This volume is the first of a series. As such, it is the first step in building a tradition. We have tried to select original essays of interest and continuing value . . . rather than to attempt a review of the published work of the previous year. The 10 to 12 articles selected for this and each of our subsequent volumes are designed to build an open-ended reference library for those who are serious about personality and social psychology. Like Bacon, I feel that some are meant to be tasted, some to be chewed, and many (if not most) to be digested over time. As the series grows, it will reflect where we are going—even more than where we have been; and, with the publication of each volume, I trust you will agree that we will share a greater insight into where we now are.

—Ladd Wheeler
Rochester, New York

1

Motive Dispositions

THE MERITS OF OPERANT AND RESPONDENT MEASURES

DAVID C. McCLELLAND

David C. McClelland is Professor of Psychology at Harvard University. His research interests are in human motivation and particularly power motivation. He is the author of *Power: The Inner Experience* (Irvington).

So-called projective tests have been attacked as scientifically valueless ever since psychologists first began using them. The case against them usually runs as follows: Their split-half and test-retest reliabilities are low. Therefore they provide unreliable measures. If the measures are unreliable, they cannot be valid. Therefore, validity coefficients sometimes obtained with them must be due to chance or some other variables. In any case, the validity coefficients often are not obtained again when they study is repeated which is only to be expected if the measures are unreliable. This is especially true when projective measures are correlated with school performance, which is the type of performance most easily and most often obtained by psychologists to ascertain the validity of any psychological characteristic. These arguments have not gone unanswered (McClelland, 1951, 1958, 1966, 1971a, 1972; Atkinson, 1960; deCharms et al., 1955; deCharms and Muir, 1978), but the charges

AUTHOR'S NOTE: I am very grateful for comments on an earlier draft of this article by J. W. Atkinson, Richard deCharms, Heinz Heckhausen, Julius Kuhl, Dan McAdams, Bernard Weiner, and David Winter.

have been reiterated in recent years with increasing vigor (Klinger, 1966; Scott & Johnson, 1972; Entwisle, 1972). To such authors it is a matter of amazement that psychologists continue to use projective test measures when the case against them is so obvious and watertight. Entwisle (1972, p. 390) quotes with approval Jensen's statement that "A satisfactory explanation of the whole amazing phenomenon is a task for future historians of psychology and will probably have to wait upon greater knowledge of the psychology of credulity than we now possess."

These authors usually also endorse personality measures obtained from structured personality questionnaires because they are internally consistent and have high test-retest reliability. Furthermore, some scores from so-called objective tests regularly correlate with school performance, the ultimate test of validity to many psychometricians. To be sure, Mischel (1968) has complained that very reliable objective personality measures do not seem to do a very good job of predicting what people will do, but the defenders of scientific virtue simply nod their heads and admit that maybe we do not have and never will have a very good science of personality anyway.

Is it as bad as all that? Is it possible that I and others continuing to use projective tests are as credulous as Jensen makes us out to be? In view of their arguments, how can I go on working with measures like *n* Achievement or *n* power (McClelland, 1975)? I thought for years that the case for the use of a measure like *n* Achievement had been so carefully and painstakingly made that psychologists generally would understand it in time and see that the attacks made on it often contained elementary errors. But I have decided that most psychologists are busy and do not have the time to follow an argument that is a little more subtle than the one on which the attacks are based. All they know is that there have been attacks and they sound reasonable. And the more often they are repeated, the more people believe there must be something to them. So it has seemed desirable to state the case once again as carefully as possible for the so-called projective measures and, in the process of doing so, to compare their merits with those of questionnaire measures of personality.

CONSTRUCT VALIDITY

Let us begin by straightening out our terminology. I came to the study of personality as an experimental psychologist interested in a general theory of behavior. This had important consequences as we shall see in a moment for what validity meant to me. It also explains why internal consistency or test-retest correlations did not interest me much. They simply are not reported in the *Journal of Experimental Psychology* and, not being trained as a psychometrician, I had not been indoctrinated in the belief that a measure has to show high reliability before it can be used. It also led me to seek a more general term than "projective test" to describe the measures we were obtaining. For "projective" refers to a theory that people are projecting their wishes into fantasy. The fact is that we were obtaining a sample of a person's behavior or thoughts in a standardized situation. I favor referring to these thought samples as operant because in Skinner's sense it is not possible to identify the exact stimulus that elicits them. More generally speaking they are responses that the subject generates spontaneously. Not the stimulus or the response or the instructional set is strictly controlled by the experimenter.

By way of contrast, *respondent* measures often specify the stimulus, the response, and the instructional set. In a typical personality test, subjects are presented with a specific stimulus (e.g., a statement like "I work like a slave . . ."); they are limited to agreeing or disagreeing with it to a greater or lesser degree; and they are set to give an overall evaluation of their behavior in a specified area. In the Thematic Aptitude test (TAT) on the other hand, the stimulus is vague, the subject can write about anything at all, and typically does not know exactly what the experimenter is getting at. That is, subjects are not being asked to conceptualize or make judgments about their behavior. There are all sorts of gradations between these two extremes, which I have discussed more fully elsewhere (McClelland, 1951). For example, some degree of stimulus control is exercised in the TAT by varying the cue characteristics of pictures. And questionnaires can ask for factual reports of behaviors, such as whether the subject jogs regularly, which can be considered operant and which do not call for the set of self-evaluation. But for the sake of simplicity I will use the terms *operant* and *respondent* from here on to refer to extreme differences in the degree of control the experimenter exercises in obtaining responses from a subject, as illustrated by the contrast between a TAT and a self-evaluative questionnaire.

What impressed me as long ago as 1951 is that operant and respondent measures generally do not correlate with each other and therefore should provide independent estimates of different aspects of personality—even when they purport to be related to the same theme.

To put it as simply as possible, the *n* Achievement measure obtained by coding operant thoughts does not correlate with what I prefer to call *v* (for value) Achievement (deCharms et al., 1955; Atkinson & Litwin, 1960) obtained from respondent questionnaires like the Edwards (1957) Personal Preference Schedule or the Jackson (1966) Personality Research Form or the Mehrabian (1969) measure of achieving tendency. This lack of correlation bothers a lot of people and they have used it as an argument that therefore, since the *v* Achievement measures are more reliable, this proves that the *n* Achievement is not valid. To me, it demonstrated that these measures get at different aspects of personality—*n* Achievement at operant trends I called motives and *v* Achievement at values which I called schemas (McClelland, 1951). Both types of variables are necessary to predict behavior (McClelland, 1979) and nothing but confusion results from insisting these two kinds of measures are assaying the same motive disposition and therefore should correlate highly.

There are many types of operants which I would classify as stylistic traits (McClelland, 1951). In a time when critics are wondering if there is anything consistent about personality (Mischel, 1968), it is worth recalling that some of them—like verbal productivity in writing stories (Entwisle, 1972), length of speech bursts (Takala, 1977), and expressive movements (Allport and Vernon, 1931)—are quite consistent. They deserve a lot more attention in personality study, but to discuss them here would be to expand this review out of shape. So let us focus on motive measures obtained from operant thoughts, although not all measures obtained from operant thoughts reflect motives (see Loevinger, 1966; Stewart, Note 1).

How do you decide that something is a measure of a motive? The answer is clear in terms of the general theory of behavior. Ever since Melton's (1952) synthesis of earlier work, it has been generally agreed that motives drive, direct, and select behavior. Hunger in a rat (or any other animal) makes it more active *(drives)*, focuses its attention on some stimuli more than others *(directs)*, and facilitates learning a maze to get food. That is, if there is food reward, its satisfaction of the hunger drive *selects* out responses that lead to the food reward. Early work with the *n* Achievement score, and later work with the other motive scores, was aimed at demonstrating that the scores were measures of a motive because people who scored high in fact behaved in these three ways as if they were more motivated. Table 1, reproduced from an earlier publication (McClelland, 1971a), illustrates how the *n* Achievement

TABLE I
Validity of Fantasy and Self-Report Measures of *n* Achievement in Predicting Behavior

Motive Function	*Fantasy Measure*	*Self-Report Measure*
A. Directing	Percentage taking moderate risks in a ring toss game	
High n Ach	61%	36%
Low n Ach	36	62
	$p = .04$	$p = .02$
B. Driving	Percentage persisting longer than average in a final examination	
High n Ach	60	42
Low n Ach	32	55
	$p = .03$	n.s.
C. Selecting	Percentage above average in final examination grade	
High n Ach	64	58
Low n Ach	32	46
	$p = .02$	n.s.

SOURCE: After Atkinson and Litwin 1960.

measure satisfies these three criteria of whether a motive is operating. Subjects scoring high in *n* Achievement persist longer, focus on moderate, challenging goals, and demonstrate that they have learned more in the course. These findings have been chosen for presentation because they make the main points of the argument in the context of a single experiment, not because they are the only or the best evidence of the functional characteristics of the achievement motive. For example, the *directing* function of *n* Achievement has been demonstrated in tachistoscopic recognition of achievement-related words (McClelland & Liberman, 1949), its *driving* function in the larger number of entrepreneurial acts it engenders (Warner et al., cited in McClelland & Winter, 1969), and its *selecting* function in the faster learning it promotes of moderately difficult materials (McClelland et al., 1953).

Table 1 also shows the results for another purported measure of *n* Achievement derived from a respondent questionnaire instrument, the Edwards Personal Preference Inventory. The *v* Achievement measure fails all three tests for the presence of a motive and therefore, if the logic of general behavior theory is correct, it cannot be considered a measure of a motive. In my terminology it is a measure of something else—the conscious value placed on achievement—a part of the self-picture. Years ago deCharms et al. (1955) demonstrated that *v* Achievement has different behavior correlates than *n* Achievement. For example, people high in *v* Achievement are more influenced by expert opinion than those low in *v* Achievement, whereas expert opinion makes no difference to people high and low in *n* Achievement. Why has it been so difficult to get psychologists to acknowledge that these are two different and generally *uncorrelated* measures of different aspects of personality though they often bear the same name? For that matter, why do psychometricians so preoccupied with the issue of the reliability of their measures repeatedly fail to take the elementary step of demonstrating that the respondent measures are in fact measures of *motives* rather than attitudes or values?

For obvious reasons psychologists do not like coding stories written to pictures; it is expensive, time-consuming, and difficult to train scorers to high coding reliability.[1] Therefore, many attempts have been made, including several in my laboratory, to find an objectively scorable instrument which will correlate highly enough with the *n* Achievement measure to qualify as a substitute for it. So far as I know, all of these efforts have failed, for the very good reason, I tend to believe, that operant and respondent measures generally tap theoretically distinct aspects of personality.

The finding shown in Table 1 of a relation between *n* Achievement and course grades will certainly not go unchallenged by the critics of operant thought measures. They have gone to great lengths to assemble published and unpublished data on the relationships between *n* Achievement and performance both at molar and task levels (Klinger, 1966; Entwisle, 1972). Their conclusion is that while there are some positive relationships, there are at least an equal number of nonsignificant relationships, which is what one would expect according to Entwisle, if the *n* Achievement measure is unreliable. "Recent studies, carefully done and more extensive in scope than earlier studies, yield few positive relationships between need achievement and other variables" (Entwisle, 1972). The examples she gives to back up this statement are drawn largely from studies of classroom performance and behavior, the favorite proving ground for validity among psychometricians. So is it not a biased selection of the facts to suggest, as Table 1 does, that *n* Achievement leads to superior performance in the classroom?

Yes it is, and I included the information on examination performance in Table 1 partly for simplicity of exposition but chiefly to make a point about the relation of *n* Achievement to performance in the most dramatic way possible. For the fact is that both friends and foes of the *n* Achievement measure have repeatedly made a basic theoretical error in assuming that there should be a direct simple relationship between *n* Achievement and performance—in all kinds of situations. Ever since the publication of *The achievement motive* (McClelland et al., 1953), it has been known that *n* Achievement has a very variable and probably overall insignificant relationship to school performance. It amazes me that so much energy has gone into proving this point over and over again. And ever since the publication of Atkinson's (1957) ground-breaking theoretical contribution on risk-taking, we have known the reason why. What his theory states in a nutshell is that moderate risk-taking (or challenge in the performance situation) is the chief incentive for people high in *n* Achievement. They will work harder than other people when the probability of success is moderate *but not if the probability of success is very high or very low.*

In other words, the incentive that people high in *n* Achievement are concerned about (moderate probability of doing better) has to be present in the situation in order for them to perform better than those low in *n* Achievement. Some findings reported by French (1955) a generation ago illustrate the point in another way. She observed how hard people high and low in *n* Achievement worked when the incentives provided in the situation differed. For one group the incentive was "to perform as well as possible," for another the incentive was to help the experimenter, and for a third group the incentive was to be excused from work. She found that *only* under the first incentive condition (doing better) did those high in *n* Achievement perform significantly better than those low in *n* Achievement. And when the incentive was time off from work, those with low *n* Achievement actually performed somewhat better. An elementary principle of general behavior theory is that an incentive or reward must be appropriate to the drive if the drive is to facilitate the acquisition of responses relative to satisfying it. A hungry rat will not learn to run a maze faster unless food is provided as the reward. If Entwisle and Jensen are incredulous over my simplemindedness, may I confess to a similar incredulity over their apparent ignorance of so elementary a principle?

So it is impossible to determine whether there should be a relation between *n* Achievement and performance in most of the instances which Klinger (1966) and Entwisle (1972) have laboriously collected because we do not know what the incentive conditions were. We do know that