

PERFORMANCE
MEASUREMENT
AND
THEORY

EDITED BY

Frank Landy

The Pennsylvania State University

Sheldon Zedeck

University of California, Berkeley

Jeanette Cleveland

Baruch College, C.U.N.Y.

TECHNICAL EDITOR

Ann Landy



1983

LAWRENCE ERLBAUM ASSOCIATES PUBLISHERS
Hillsdale, New Jersey

London

COMMENTS

In Defense of Surface Structure

Richard A. Shweder
University of Chicago

Some ideas are so good they deserve to be true, but (alas) are not. One of those deserving, but misleading, ideas is the idea that mind and personality consist of abstract underlying structures (e.g., "reversible operations," "anxiety-proneness") that impose form on any concrete task, problem, or situation that happens to come along. The difficulty with the idea of "abstract structure" is the implication that one can be indifferent to content and still make sense of actual cognitive, affective, and interpersonal functioning. What many cognitive scientists and personality researchers have discovered is that one cannot be indifferent to whether (for example) the intellectual task is conservation of *number* or conservation of *liquid quantity* or whether the affect-eliciting stimulus is a *job interview* or a *dentist's drill* (e.g., Brainerd, 1978; Flavell, 1980; Mischel, 1968, 1973; Mischel and Peake, 1981; Moos, 1969). Whether or not a particular abstract structure (reversible operations, anxiety) gets applied depends on the task.

Allport (1960) buys in to the idea of "abstract structure." In his classic discussion of personality traits, he notes that "the stimulus is not the crucial determinant in behavior that expresses personality; the trait itself is decisive. Once formed, a trait seems to have the capacity of directing responses to stimuli into characteristic channels [p. 132]."

Piaget shares Allport's Platonic image: "Object constancy," "reversible operations," "the predicate calculus" are all abstract representations of mental capacities that are supposed to direct cognitive responses to particular tasks into characteristic and general channels of functioning.

This essay is a brief defense of content, the task, the stimulus, "surface structure." In contrast to Allport and Piaget I argue that one cannot afford to be indifferent to content or to "apparently" irrelevant details of a task or situation,

especially if one's goal is performance measurement or the assessment of actual functioning. Perhaps those who study visual perception can be indifferent to content. The laws of visual perception may be the same whether you are viewing a boxing match or staring at a cow. But how people think, feel, and act is not packaged or stabilized at the abstract level of (for example) "anxiety-ridden" functioning, "dependent" functioning, or "formal operational" functioning. The person who starts biting his or her nails, sweating, and fidgeting during an interview is not typically the person who feels tense and jittery when speaking before a group. The person who is unflappable when trapped in an elevator is no less likely than a blanching elevator mate to "go to pieces" when he or she finds him- or herself unable to solve an important problem at work (Mischel, 1968, 1973; Mischel and Peake, 1981; Van Heck, 1981). Attention to detail—attention to the distinguishing and special features of particular performance sites—is where the action is.

Consider, for example, intellectual functioning. Recent work suggests that by varying the content of a task it is possible to elicit preoperational thinking from an adult (Wason, 1969; Wason & Johnson-Laird, 1972) or formal operational thinking from a four year old (Macnamara, Baker, & Olson, 1976); and that the person who functions at a formal operational level on one problem is not typically the same person who functions at a formal operational level on a second problem. Indeed, to cite one example, Roberge and Flexer (1979) discover that performance on formal operational tests for propositional logic and combinatorial thinking correlate a mere $-.07$ (for eighth graders) and $.17$ (for adults). They are joined by others in their conclusion (Roberge & Flexer, 1979) that:

if, as Piaget and others hypothesize, formal operational thinking reflects an organized structure of second-order operations, one would expect to find a basic consistency in performance across tasks, particularly for groups of subjects who are presumably well-established in the formal operational stage. However, the minimal association between scores on the tests . . . provides little evidence of the functional interdependence among these logical operations that is claimed by proponents of the structural whole (structure d'ensemble) model of formal operations [p. 482]. (Also see Brainerd, 1978.)

Intellectual functioning is often task specific and, thus, manipulable. Wason and Johnson-Laird (1972; also Wason, 1969), for example, present subjects with a formal operation "hypothetical contradiction task." Imagine the following task: You are presented with a series of cards. You are told that each card has a number on one side and a letter on the other. You have been hired to inspect cards to make sure that a certain rule for producing cards has not been violated. That rule is: *If there is a vowel on one side of the card there must be an odd number on the other side.* Four cards are on the table before you: K, A, 7, 8.

Remember, each of these four cards has a letter on one side and a number on the other. Now, which and only which of the four cards *must* be inspected to make sure the rule is not violated? Answer this question before reading any further!

Wason and Johnson-Laird discover that fewer than 10% of college-educated adults draw the proper deduction from the rule. "If there is a vowel on one side there must be an odd number on the other side" (if p then q) implies "if there is not an odd number on one side then there must not be a vowel on the other side" (if not q then not p); hence, inspect the A and the 8! Furthermore, many subjects have difficulty recognizing the source of their error. Thus, they recognize that an "A" card with an "8" on the back is a rule violation, but they fail to mentally rotate that card. They do not recognize that an "A" card with an "8" on the back is the same as an "8" card with an "A" on the back. Hence, also inspect the "8." In other words, subjects seem to lack "reversible operations"; they function at a preoperational level. Yet, when Wason and Johnson-Laird introduce slight changes in the content (but not the abstract form) of the task, a majority of their subjects all of a sudden perform at a formal operational level. Wason and Johnson-Laird (1972) conclude that "formal operational thought is less general than Piaget supposed [p. 193]." More importantly, Wason and Johnson-Laird demonstrate that details of the content about which you think can be decisive for how you cognitively function.

A similar point is made by Van Heck (1981) with regard to affective functioning (also see Shweder, 1979, 1981). Van Heck examines the abstract personality structure known as "proneness to anxiety." Anxiety is assessed utilizing self-report, observational, and physiological measures. Van Heck also analyzes details of the performance sites in which anxiety might be observed. He discovers that his subjects perceive anxiety-provoking situations as coming in (at least) four kinds: (1) threats to interpersonal status (e.g., a job interview); (2) inanimate environmental threats (e.g., sitting on an airplane just before takeoff); (3) threats of punishment (e.g., breaking something that does not belong to you); and (4) threats of pain (e.g., the dentist's drill approaches your mouth). Van Heck discovers that those who are more anxious than others in one of these types of situations are not typically more anxious than others in the other situations. The content of the situation matters: It is the decisive determinant of who is more or less anxious.

Van Heck's observational data are quite illuminating. Subjects are observed in diverse anxiety-provoking situations (oral exams, public speaking, threatened with electrical shock, confronted with an unsolvable problem, etc). Behavior in these situations is coded for diverse response modes, all theoretically related to anxiety (restless hand movements, tense laughing, downward gaze, posture, etc.). Van Heck discovers that 0% (I repeat, 0%) of the variance of observed behavior is accounted for by stable individual differences in the (purported) abstract structure "proneness to anxiety." The way people line up on anxiety in an oral exam tells you little about how those same people line up on anxiety when

they get "stuck" solving a problem. Only 1% of the variance of observed behavior is accounted for by stable situational differences. In other words, how situations line up in terms of being anxiety provoking for one person tells you little about how those same situations line up in terms of being anxiety provoking for the next person. A job interview gets on some people's nerves more than a dentist's drill, and vice versa for other people. Notably, a whopping 33% of the variance is accounted for by stable response modes; in other words, anxiety is more likely to get expressed in some ways (biting lips) than other ways (nervous laughter) and there is some tendency for this "hierarchy of response mode likelihood" to be stable across persons and situations.

Van Heck's findings are not idiosyncratic. Mischel and Peake (1981) examined the purported abstract structure "conscientiousness" among college students. Utilizing reliable, aggregated measures of such variables as "room neatness," "assignment neatness," "class attendance," and so on, and correcting their measures for attenuation, Mischel and Peake discovered average cross-situational consistency coefficients of .20. The mundane world of cognitive, affective, and interpersonal functioning does not seem to instantiate the supposed directive influence of underlying abstract structures (also see Mischel, 1968).

Why have we been so indifferent to content in the study of human performance? Perhaps one answer is that we all too readily assume that to do "science" is to search for deep structures and general processes. The "periodic table of elements" model of science has a firm grip on the brains of many social scientists. Perhaps a second answer is that occasionally the search for deep structures and general processes is successful. A third possible answer is suggested by the "systematic-distortion hypothesis" (D'Andrade, 1965, 1974; Shweder, 1975, 1977a, 1977b, 1981; Shweder & D'Andrade, 1979, 1980), with special reference to our understanding of individual differences in social behavior. Borman (this volume) provides an excellent overview of this line of research. The basic idea is that many of the most popular social-science procedures for gathering data on personality differences (summary ratings by peers or observers) stimulate the rater to report conceptually associated memory items as though they actually cooccurred.

For example, I suspect systematic-distortion processes may be at work in the summary ratings by teacher-observers reported in Moskowitz (1978). Focusing on the abstract structures "dependency" and "dominance," Moskowitz assesses the performance of children in a preschool. Over an eight-week period of time, four hours of on-line observational data are collected on each of 56 children. "Dependency" is assessed for each of five response modes or content areas: seeking help, seeking supervision, being near, touching, and seeking recognition. "Dominance" is assessed for each of five response modes or content areas: commanding, threatening, directing, displacing, and making suggestions. Classroom teachers (the children's "supervisors," so to speak) were asked to make summary ratings on the children for each response mode or

content area. What is striking is that in the teachers' summary ratings the response modes or content areas seem almost mutually substitutable, each an expression of the (purported) abstract motivational structure "dependency" or "dominance." For the summary ratings of teachers the average intratrait consistency coefficient is .50 for dependency and .57 for dominance. Thus, for example, "seeking help" and "seeking supervision" are correlated .74 in the teachers' ratings. "Threatening others" and "displacing others" are correlated .94. However, in the direct observational data "seeking help" and "seeking supervision" correlate a mere .14 whereas "threatening others" and "displacing others" correlate $-.03$. For the on-line observational data the average intratrait consistency coefficient is .04 for dependency and .25 for dominance. Remember, this is, observational data aggregated over eight weeks of observation.

Summary-rating tasks seem to direct the attention of observers to the conceptual relatedness of rating items; theoretically associated items get reported as though they went together despite observational evidence to the contrary. As D'Andrade (1965) remarked: One of the hazards of science is the ease with which we confuse "propositions about language" with "propositions about the world [p. 215]." The search for abstract personality structures has been fueled by that confusion.

When we do look at the world of individual differences in social behavior using reliable, "on-line" observational teachings, what we discover is that "generalizations decay" (Cronbach, 1975). And, they do not decay over time alone. They decay across minor changes in task, context, or response mode. If one wanted to know whether some child is going to spend time "displacing" other children one would not want to examine the child's "threat" behavior. The "wise" stay as close to the situation and response mode of interest as possible. Indeed, individual differences in many domains (e.g., anxiety) are so dynamic and unstable that individual differences destabilize as people become more familiar with the particular performance site. It is less familiar situations that elicit more consistent behavior (Van Heck, 1981).

Borman's excellent contribution to this volume speaks for itself; I have only a few remarks to make. The systematic-distortion hypothesis is sometimes cast as a problem in memory retrieval and sometimes as a problem in information integration. Borman is sensitive to the distinction. Unassisted, our everyday information-integration skills can be quite deficient. For example, imagine you know that "John has self-esteem." Which is the better inference: (1) therefore, John probably is not a leader; or (2) therefore, John probably is a leader? Most observers choose the second inference. Yet when interrogated most observers believe that approximately 50% of the people they know have self-esteem whereas only 15% are leaders. Because the observers' own base-rate estimates suggest that "most people with self-esteem are not leaders" it is the first inference just cited that should be endorsed. Integrating our inferential judgments with what we

think we know is not a trivial task. More work needs to be done on the cognitive processes (memory, information integration) that produce systematic distortion.

Borman identifies factors leading to the stability of individual differences. For example, he notes that in many work settings the lack of variation in performance sites favors consistency in behavior. The point is a sound one. Note, however, that the lack of variation in a performance site may be offset by the "familiarity" effect previously noted. Individual difference rankings sometimes destabilize just because a setting remains the same. Under familiar conditions behavior spontaneously varies.

Borman also notes that individual differences in skills are more stable than individual differences in motivation or interpersonal behavior. Again, this is a nice point, although I would emphasize that not all skilled performances are merely the expressions of individual talent and training. Skilled performances vary in their dependency on props, interpersonal relationships, and interpersonal coordination. I would bet that "underdogs" are less successful in tennis or squash than in football or basketball (where "upsets" are probably more likely).

Finally, I'm less inclined than Borman to treat cognitive functioning as a special case. I.Q., or "brightness," may well be more general than most "traits" of social behavior (e.g., dependency). However, I.Q. is only one part of the story on mind. In many areas of cognitive functioning, individual differences can be quite task specific, as my brief excursion into formal operational thinking suggested. Freud is well-known for his slogan: Where there is Id let there be Ego. Perhaps less well-known is the slogan: Where there is structure let there be content!

REFERENCES

- Allport, G. W. *Personality and social encounter*. Boston: Beacon Press, 1960.
- Brainerd, C. J. *Piaget's theory of intelligence*. Englewood Cliffs, N.J.: Prentice-Hall, 1978.
- Cronbach, L. J. Beyond the two scientific disciplines of scientific psychology. *American Psychologist*, 1975, 30, 116-127.
- D'Andrade, R. G. Trait psychology and componential analysis. *American Anthropologist*, 1965, 67, 215-228.
- D'Andrade, R. G. Memory and the assessment of behavior. In T. Blalock (Ed.), *Measurement in the social sciences*. Chicago: Aldine-Atherton, 1974.
- Flavell, J. H. *Structures, stages and sequences in cognitive development*. Paper presented at 1980 Minnesota Symposium on Child Psychology. (Available from J. H. Flavell, Department of Psychology, Stanford University, Stanford, California 94305.)
- Macnamara, J., Baker, E., and Olson, C. L. Four-year-olds' understanding of *pretend, forget and know*: Evidence for propositional operations. *Child Development*, 1976, 47, 62-70.
- Mischel, W. *Personality and assessment*. New York: Wiley, 1968.
- Mischel, W. Towards a cognitive social learning reconceptualization of personality. *Psychological Review*, 1973, 80, 252-283.
- Mischel, W., and Peake, P. In search of consistency: Measure for measure. In M. P. Zanna, E. T. Higgins, and C. P. Herman (Eds.), *Consistency in social behavior: The Ontario symposium of personality and social psychology* (Vol. 2). Hillsdale, N.J.: Lawrence Erlbaum Associates, 1981.

- Moos, R. H. Sources of variance in responses to questionnaires and in behavior. *Journal of Abnormal Psychology*, 1969, 74, 405-412.
- Moskowitz, D. S. Analyzing personality constructs for internal consistency and transsituational generality. Unpublished doctoral dissertation, University of Connecticut, 1978. *Dissertation Abstracts International*, 1979, 40, 4328 (University microfilms #7914176).
- Roberge, J. J., & Flexer, B. K. Further examination of formal operational reasoning abilities. *Child Development*, 1979, 50, 478-484.
- Shweder, R. A. How relevant is an individual difference theory of personality? *Journal of Personality*, 1975, 43, 455-484.
- Shweder, R. A. Illusory correlation and the M.M.P.I. controversy. *Journal of Consulting and Clinical Psychology*, 1977, 45, 917-924. (a)
- Shweder, R. A. Likeness and likelihood in everyday thought: Magical thinking in judgments about personality. *Current Anthropology*, 1977, 18, 637-648. (b)
- Shweder, R. A. Rethinking culture and personality theory Part I: A critical examination of two classical postulates. *Ethos*, 1979, 7, 255-278.
- Shweder, R. A. Fact and artifact in trait perception: The systematic distortion hypothesis. In B. A. Maher & W. B. Maher (Eds.), *Progress in experimental personality research* (Vol. 11). New York: Academic Press, 1981.
- Shweder, R. A., & D'Andrade, R. G. Accurate reflection or systematic distortion? A reply to Block, Weiss and Thorne. *Journal of Personality and Social Psychology*, 1979, 37, 1075-1084.
- Shweder, R. A., & D'Andrade, R. G. The systematic distortion hypothesis. In R. A. Shweder (Ed.), *Fallible judgment in behavioral research. New directions for methodology of social and behavioral science*, No. 4. San Francisco: Jossey-Bass, 1980.
- Van Heck, G. L. M. *Anxiety: The profile of a trait*. Unpublished doctoral thesis, Department of Psychology, Tilberg University, The Netherlands, 1981. (Available from G. L. M. Van Heck, Department of Psychology, Tilberg University, P.O. Box 90153, 5000 Tilberg, The Netherlands.)
- Wason, P. C. Regression in reasoning? *British Journal of Psychology*, 1969, 60, 471-480.
- Wason, P. C., & Johnson-Laird, P. N. *Psychology of reasoning*. London: B. T. Batsford, 1972.

