#99

In P. E. Meehl (1973) *Psychodiagnosis: Selected papers* (pp. 225-302, Chapter 13). Minneapolis: University of Minnesota Press.

13
Why I Do Not Attend Case Conferences

I HAVE FOR MANY YEARS been accustomed to the social fact that colleagues and students find some of my beliefs and attitudes paradoxical (some would, perhaps, use the stronger word *contradictory*). I flatter myself that this paradoxicality arises primarily because my views (like the world) are complex and cannot be neatly subsumed under some simple-minded undergraduate rubric (e.g., behavioristic, Freudian, actuarial, positivist, hereditarian). I find, for example, that psychologists who visit Minneapolis for the first time and drop in for a chat with me generally show clinical signs of mild psychic shock when they find a couch in my office and a picture of Sigmund Freud on the wall. Apparently one is not supposed to think or practice psychoanalytically if he understands something about philosophy of science, thinks that genes are important for psychology, knows how to take a partial derivative, enjoys and esteems Albert Ellis, or is interested in optimizing the prediction of behavior by the use of actuarial methods! I maintain that there is no unresolvable conflict between these things, but do not propose to argue that position here.

On the local scene, one manifestation of this puzzlement has come frequently to my attention and, given its nature, I think it likely that for each time I hear the question there are numerous other occasions when it is raised. In substance, the puzzle–sometimes complaint–among our graduate students goes like this: "Dr. Meehl sees patients on the campus

and at the Nicollet Clinic, averaging, so we are told, around a dozen hours a week of psychotherapy. With the exception of a short period when he was APA president, he has been continuously engaged in the practice of psychotherapy for almost thirty years. It is well known that he not only thinks it important for a psychologist to work as a responsible professional with real-life clinical problems but, further, considers the purely 'theoretical' personality research of academic psychologists to be usually naive and unrealistic when the researcher is not a seasoned, practicing clinician. When he taught the introductory assessment course, the lectures were about evenly divided between rather abstract theoretical and methodological content (such as 'What is the nature of a phenotypic trait, considered as a class of related dispositions?' 'What precisely do we mean by the phrase *disease entity*?' 'What is *specific etiology*?') and practical, down-to-earth material (such as 'How do you handle a patient's questions about yourself?' 'What do you do with the patient who in the initial interview sits passively expecting you to cross-examine him?' 'How do you assess the severity of a depression, especially with respect to suicidal potential?' 'How do you tell the difference between an acting-out neurotic and a true psychopath?'). He took the trouble to become a (non-grandfathered) diplomate of ABPP although in his academic position this had little advantage either of economics or of status. When he was chairman of the Psychology Department he had a policy of not hiring faculty to teach courses in the clinical and personality area unless they were practitioners and either had the ABPP diploma or intended to get it. He has been an (unsuccessful) advocate of a special doctorate in clinical psychology, the Ps.D., which would dispense with some of the medieval academic requirements for the Ph.D. degree and would permit a much more intensive and diversified clinical training for persons aiming at full-time work as practitioners in the profession. Meehl lists himself in the Yellow Section of the phone book and is a member of such outfits as the American Academy of Psychotherapists, the American Academy of Psychoanalysis, and the Institute for Advanced Study in Rational Psychotherapy. On all these counts, it seems evident that Meehl is 'clinically oriented,' that his expressed views about the importance of professional practice are sincere rather than pro forma. It is there-fore puzzling to us students, and disappointing to us after having been stimulated by him as a lecturer, to find that he almost never shows up in the clinical settings where we take our clerkship and

internship. We never see Dr. Meehl at a case conference. Why is this?"

This understandable puzzlement was the precipitating cause of my writing the present paper, partly because it becomes tiresome to explain this mystery repeatedly to baffled, well-meaning students, but also because responding to the puzzlement provides an occasion for some catharsis and, I hope, for making a constructive contribution to the field. Accordingly the first portion of the paper will be highly critical and aggressively polemic. (If you want to shake people up, you have to raise a little hell.) The second part, while not claiming grandiosely to offer a definitive solution to the problem, proposes some directions of thinking and "experimenting" that might lead to a significant improvement over current conditions.

The main reason I rarely show up at case conferences is easily stated: The intellectual level is so low that I find them boring, sometimes even offensive. Why the level of a psychiatric case conference is usually so mediocre, by contrast with conferences in internal medicine or neurology—both of which I have usually found stimulating and illuminating—is not known, and it is a topic worthy of research. I do not believe my attitude is as unusual as it may seem. I think I am merely more honest than most clinical psychologists about admitting my reaction. Witness the fact that the staff conferences in the Medical School where I work are typically attended by only a minority of the faculty—usually those who *must* be there as part of their paid responsibility, or who have some other special reason (such as invitation) for attending a particular one. If the professional faculty found them worthwhile, they wouldn't be so reluctant to spend their time that way. Pending adequate research on "What's the matter with the typical case conference," I present herewith some clinical impressions by way of explanation, and some constructive suggestions for improvement. My impressionistic list of explanations constitutes the "destructive criticism" portion of this paper.

Part I: What's Wrong?

1. *Buddy-buddy syndrome*. In one respect the clinical case conference is no different from other academic group phenomena such as committee meetings, in that many intelligent, educated, sane, rational persons seem to undergo a kind of intellectual deterioration when they gather around a table in one room. The cognitive degradation and feck-

less vocalization characteristic of committees are too well known to require comment. Somehow the group situation brings out the worst in many people, and results in an intellectual functioning that is at the lowest common denominator, which in clinical psychology and psychiatry is likely to be pretty low.

2. "*All evidence is equally good*." This absurd idea perhaps arises from the "groupy," affiliative tendency of behavioral scientists in "soft" fields like clinical, counseling, personality, and social psychology. It seems that there are many professionals for whom committee work and conferences represent part of their social, intellectual, and erotic life. If you take that "groupy" attitude, you tend to have a sort of mush-headed approach which says that everybody in the room has something to contribute (absurd on the face of it, since most persons don't usually have anything worthwhile to contribute about anything, especially if it's the least bit complicated). In order to maintain the fiction that everybody's ideas are worthwhile, it is necessary to lower the standards for what is evidential. As a result, a casual anecdote about one's senile uncle as remembered from childhood is given the same group interest and intellectual respect that is accorded to the citation of a high-quality experimental or field-actuarial study. Or a casual impression found in the nurses' notes is given the same weight as the patient's MMPI code. Nobody would be prepared to defend this rationally in a seminar on research methods, but we put up with it in our psychiatric case conferences.

3. *Reward everything—gold and garbage alike*. The tradition of exaggerated tenderness in psychiatry and psychology reflects our "therapeutic attitude" and contrasts with that of scholars in fields like philosophy or law, where a dumb argument is *called* a dumb argument, and he who makes a dumb argument can expect to be slapped down by his peers. Nobody ever gives anybody negative reinforcement in a psychiatric case conference. (Try it once—you will be heard with horror and disbelief.) The most inane remark is received with joy and open arms as part of the groupthink process. Consequently the educational function, for either staff or students, is prevented from getting off the ground. Any psychologist should know that part of the process of training or educating is to administer differential reinforcement for good versus bad, effective versus ineffective, correct versus incorrect behaviors. If *all* behavior is rewarded by friendly attention and nobody is ever non-

reinforced (let alone punished!) for talking foolishly, it is unlikely that significant educational growth will take place.

A corollary of the "reward everything" policy with respect to evidence and arguments is a substantive absurdity, namely, everyone is right—or at least, nobody is *wrong*. The group impulse toward a radical democratization of qualifications and opinions leads almost to denying the Law of Noncontradiction. A nice quotation from the statistician M. G. Kendall is apposite: "A friend of mine once remarked to me that if some people asserted that the earth rotated from East to West and others that it rotated from West to East, there would always be a few well-meaning citizens to suggest that perhaps there was something to be said for both sides and that maybe it did a little of one and a little of the other; or that the truth probably lay between the extremes and perhaps it did not rotate at all" (Kendall, 1949, p. 115) .

4. *Tolerance of feeble inferences* (*e.g., irrelevancies*). The ordinary rules of scientific inference, and reliance upon general principles of human development, which everybody takes for granted in a neurology staff conference, are somehow forgotten in a psychiatric case conference. This is perhaps due to the fact that the psychiatrist has had to learn to live with the sorry state of his specialty after having had training in the more scientific branches of medicine, with the result that once having learned to live this way, he assumes that the whole set of rules about how to think straight have to be junked, so that logic, statistics, experiments, scientific evidence, and so on don't apply. I have heard professionals say things in a psychiatric staff conference which I am certain they would never have said about a comparable problem in a conference room one floor below (neurology service). Example: In a case conference involving a differential diagnosis between schizophrenia and anxiety reaction in a pan-anxious patient that any well-read clinician would easily recognize as a classical case of the Hoch-Polatin "pseudoneurotic schizophrenia" syndrome (Hoch and Polatin, 1949; Meehl, 1964) the psychiatrist presiding at the conference argued that the patient was probably latently or manifestly schizophrenic. He argued thus partly because—in addition to her schizophrenic MMPI profile—she had a vivid and sustained hallucinatory experience immediately preceding her entry into the hospital. She saw a Ku Klux Klansman standing in the living room, in full regalia, eyeing her malignantly and making threatening gestures with a knife, this hallucination lasting for several

minutes. Since hallucinations of this sort are textbook symptoms of a psychotic break in ego function (reality testing), it seemed pretty clear to the presiding psychiatrist (and myself) that this would have to be considered evidence—not dispositive, but pretty strong—for our schizophrenic diagnosis as against the anxiety-neurosis alternative. At this point one of the nurses said, "I don't see why Dr. Koutsky and Dr. Meehl are laying emphasis upon this Ku Klux Klansman. After all, I remember having an imaginary companion when I was a little girl." Now suppose that this well-meaning nurse, whose remark was greeted with the usual respectful attention to "a contribution," had been attending a case conference on the neurology service. And suppose that in attempting a differential diagnosis of spinal cord tumor the presiding neurologist had offered in evidence the fact that the patient was incontinent of urine. It would never occur to this nurse to advance, as a counterargument, the fact that she used to wet her pants when she was a little girl. (If she did advance such a stupid argument on neurology, my colleague Dr. A. B. Baker—who has "standards"—would tromp on her with his hobnail boots, and she would never make *that* mistake again.) But somehow when she gets into a psychiatric case conference she undergoes a twenty-point decrement in functional IQ score, so as to forget how to distinguish between different degrees of pathology or between phenomena occurring at different developmental levels. Equating a childhood imaginary companion with an adult's experiencing a clear and persisting visual hallucination of a Ku Klux Klansman is of course just silly—but in a psychiatry case conference no one would be so tactless as to point this out.

5. *Failure to distinguish between an inclusion test and an exclusion test*: In a differential diagnosis between schizophrenia and manic-depressive psychosis, a psychology trainee argues against schizophrenia on the ground that the patient does not have delusions or hallucinations with clear sensorium. Of course this is just plain uninformed, because delusions and hallucinations are among Bleuler's "accessory" symptoms, present in some schizophrenics but not all, and they are *not* part of the indicator family that "defines" the disease (Bleuler, 1911 as reprinted 1950). Some American clinicians (not I) would hold that delusions and hallucinations with clear sensorium are so rare in uncomplicated manic depression that when *present* they could be used as a quasi-exclusion test against that diagnosis. But since many schizophrenics—not only

borderline cases of "pseudoneurotic schizophrenia" but those cases known in the present nomenclature as "schizophrenia, chronic undifferentiated" and "schizophrenia, acute episode" and "schizophrenia, simple type"—are without these particular accessory symptoms, the trainee's argument is without merit. Psychodynamically, delusions and hallucinations are among the so-called restitutional symptoms of the disorder, as contrasted with the regressive ones. Depending upon the form and stage of the disease, restitutional symptoms may or may not be in evidence. That delusions and hallucinations with unclouded sensorium are absent in many schizophrenics is not an idiosyncratic clinical opinion of mine. It is a theory found in all of the textbooks, it is in the standard nomenclature, it is in Kraepelin and Bleuler, who defined the entity "schizophrenia." There is no justification for utilizing the *absence* of these accessory symptoms as an exclusion test. Neither semantic nor empirical grounds exist for this practice. But when I point this out forcefully, the trainee looks at me as if I were a mean ogre.

6. *Failure to distinguish between mere consistency of a sign and differential weight of a sign*. Once the differential diagnosis has been narrowed to two or three nosological possibilities, it is inappropriate to cite in evidence signs or symptoms which are nondifferentiating as between them. This is so obvious a mistake that one thinks it would never happen; but some clinicians do it regularly. In distinguishing *between* a sociopathic personality, an acting-out neurotic delinquent, and a garden-variety "sociological" criminal, it is fallacious to argue that the patient was a marked underachiever or a high school dropout, in spite of high IQ, as grounds for a diagnosis of sociopathic personality, because, whereas this sign is a correlate of the sociopathic diagnosis, we have now narrowed the nosological range to three possibilities, each of which is a correlate of academic underachievement, so that this sign has lost its diagnostic relevancy at this stage of the investigation. This illustrates one of the generic features of case conferences in psychiatry, namely, the tendency to mention things that don't make any difference one way or the other. The idea seems to be that as long as something is true, or is believed to be true, or is possibly true, it is worth mentioning! In other medical specialties in order to be worth mentioning the statement must not only be true but be *differentially* relevant, i.e., it must argue for one diagnosis, outlook, or treatment, rather than another.

7. *Shift in the evidential standard, depending upon whose ox is being*

gored. A favorite tactic of case conference gamesmanship is to use a "double standard of morals" on the weight of the evidence. When you are putting your own diagnostic case, you permit indirect inferences (mediated by weak theoretical constructions and psychodynamic conclusions); then when the other fellow is making his case for a different diagnosis, you become superscientific and behavioristic, making comments like "Well, of course, all we actually know is the *behavior*." You *don't really* know "the behavior" in the sense it is usually discussed in the staff conference, since even phenotypic characterizations are almost invariably summary-type statements with a large component of sampling inference at least involved. Further, to this sampling inference we usually conjoin theory-mediated inferences, relying on extrapolations from other contexts as justification for weighting some sources of data more heavily than others. As a result this superbehaviorism is not even intellectually honest.

The opposite of this ("simpleminded") error is, of course, the failure to connect theoretical constructs with behavioral data, actual *or possible*. This is the error of the "muddleheaded." Projective tests lend themselves particularly well to this, since trends, forces, and structures that are *latent* (a perfectly legitimate metaconcept) cannot be operationally defined, hence offer unusual temptation for a muddlehead to use them without regard for any kind of corroborative evidence, direct or indirect, tight or probabilistic.

8. *Ignorance* (*or repression*) *of statistical logic*. A whole class of loosely related errors made in the clinical case conference arises from forgetting (on the part of the psychologist) or never having learned (in the case of the psychiatrist and social worker) certain elementary statistical or psychometric principles. Examples are the following:

a. *Forgetting Bayes' Theorem*. One should always keep in mind that there is a relationship between prior probability (e.g., the base rate *P* of a certain diagnosis or dynamic configuration in the particular clinic population) and the increment in probability contributed by a certain diagnostic symptom or sign. If the prior probability is extremely low, you don't get very much mileage out of a moderately strong sign or symptom. On the other hand, when the prior probability is extremely high, you get mileage out of an additional fact, but you don't really "need it much," so to speak. The considerations advanced by Meehl and Rosen (1955—reprinted here as Chapter 2) apply in a clinical

case conference just as strongly as they do in a research design involving psychometrics.

b. *Forgetting about unreliability when interpreting score changes or difference scores* (e.g., on subtests of the WAIS). Despite the mass of adverse research and psychometric theoretical criticism of the practice of overinterpreting small difference scores on unreliable subtests (which are of doubtful validity for the alleged noncognitive traits anyway!), one still hears this kind of "evidence" pressed in case conferences. Who cares whether the patient "did well on the Block Design subtest but seemed to enjoy it less than Picture Arrangement"?

c. *Reliance upon inadequate behavior samples for trait attribution*. Sometimes the inadequacy is *qualitative*, in the sense that the context in which the behavior was sampled is in some way unusual or atypical for the population or for this particular individual; more commonly, the error is simply one of believing that you can estimate the proportion of white marbles in an urn after sampling only a couple of marbles. This error is particularly serious because in addition to the numerical smallness of the samples of behavior adduced as the basis for trait attribution, we have almost no control over the conscious or unconscious selection factor that has determined which behavior chunk was noticed, was remembered, and is now reproduced for tendentious purposes. It is obvious that over a period of several hours or days of unsystematic observation, practically any human being is likely to emit at least a few behaviors which can be subsumed under almost any trait in the phenotypic or genotypic lexicon.

d. *Inadequate consideration of whether and when the* (*fact* → *fact*) *linkage is stronger or weaker than the* (*multiple-fact* → *diagnosis* → *fact*) *linkage*. It seems there are some cases in which the best way to infer to a certain fact, whether postdictive or predictive, is by relying upon its correlation with certain other relatively atomistic facts with which, from previous experience or research, the inferred fact is known to be correlated. In other cases it appears that a set of facts which qualitatively does not seem related to the fact of interest is related to it rather strongly because this first set of facts known to us converges powerfully upon a taxonomic decision (whether formal diagnosis, environmental mold, personality "type," or dynamic configuration). When that taxonomic decision has been made with high confidence, certain other individual atomistic facts or dispositions may follow with reasonably high

confidence. It is a mistake to assume, without looking into the matter, that one or the other of these approaches is "obviously" the way to proceed most powerfully. (Cf. Meehl, 1960—reprinted here as Chapter 6.)

e. *Failing to understand probability logic as applied to the single case*. This disability is apparently endemic to the psychiatric profession and strangely enough is also found among clinical psychologists in spite of their academic training in statistical reasoning. There are still tough, unsolved philosophical problems connected with the application of frequencies to individual cases. But we cannot come to grips with those problems, or arrive at a pragmatic decision policy in staff conferences, unless we have gotten beyond the blunders characteristically enunciated by clinicians who are not familiar with the literature on this subject from Lundberg (1941) and Sarbin (1942) through Meehl (1945a, 1954a, 1956a, 1956b, 1956c—reprinted here as Chapter 3, 1957—reprinted here as Chapter 4, l959a, l959b—reprinted here as Chapter 5, 1960—reprinted here as Chapter 6, Meehl and Dahlstrom, 1960) to recent contributors like Goldberg (1968, 1970), Sawyer (1966), Kleinmuntz (1968, 1969), Einhorn (1970, 1972), Pankoff and Roberts (1968), Marks and Sines (1969), Alker and Hermann (1971), Mirabile, Houck, and Glueck (1971); see also footnote 4 in Livermore, Malmquist, and Meehl, 1968 (at page 76), and footnotes 8 and 9 in Meehl, 1970b (at pp. 8-9), and references cited thereat. The vulgar error is the cliché that "We aren't dealing with groups, we are dealing with this individual case." It is doubtful that one can profitably debate this cliché in a case conference, since anyone who puts it quite this way is not educable in ten minutes. He who wishes to reform the thinking in case conferences must constantly reiterate the elementary truth that if you depart in your clinical decision making from a well-established or even moderately well-supported) empirical frequency—whether it is based upon psychometrics, life-history material, rating scales or whatever—your departure may save a particular case from being misclassified predictively or therapeutically; but that such departures are, prima facie, counterinductive, so that a decision *policy* of this kind is almost certain to have a cost that exceeds its benefits. The research evidence strongly suggests that a policy of making such departures, except very sparingly, will result in the misclassifying of other cases that would have been correctly classified had such nonactuarial departures been forbidden; it also suggests that more of this second kind of misclassification will occur than

will be compensated for by the improvement in the first kind (Meehl, 1957—reprinted here as Chapter 4). That there are occasions when you should use your head instead of the formula is perfectly clear. But which occasions they are is most emphatically *not* clear. What *is* clear on the available empirical data is that these occasions are much rarer than most clinicians suppose.

9. *Inappropriate task specification*. Nobody seems very clear about which kinds of tasks are well performed in the case conference context and which would be better performed in other ways. There are some cognitive jobs for which it seems doubtful that the case conference is suitable. I myself think that the commonest form of this mistake is the spinning out of complicated psychodynamics which are explained in terms of the life history and which in turn are used to explain the present aberrant behavior, on evidence which is neither quantitatively nor qualitatively adequate to carry out such an ambitious enterprise (assuming, as I believe, that the enterprise is sometimes feasible in the present state of psychology). Any psychologist who has practiced long-term, intensive, "uncovering" psychotherapy knows that there are psychodynamic puzzles and paradoxes which remain in his mind after listening to fifty or a hundred hours of the patient's productions. Yet this same psychotherapist may undergo a strange metamorphosis when he enters the case conference context, finding himself pronouncing (sometimes rather dogmatically) about the psychodynamics of the presented patient, on the basis of ten minutes' exposure to the patient during the conference, plus some shoddy, scanty "material" presented by the resident and social worker (based in turn upon a relatively small total time of contact with the patient and interviewing that *on the psychotherapist's own usual criteria* would be considered "superficial").

Part of the difficulty here lies in American psychiatry's emphasis upon psychodynamics at the expense of nosology. A case conference *can* be, under some circumstances, an appropriate place to clarify the nosological or taxonomic issue provided that the participants have bothered to learn some nosology, and that the clinicians mainly concerned with the patient have obtained the relevant clinical data. But since diagnosis is devalued, the prestigious thing to do is to contribute psychodynamic ideas to the conference, so we try to do that, whether or not the quality and quantity of the material available to us is adequate to such an enterprise, which it usually isn't.

10. *Asking pointless questions*. Participants in a case conference frequently ask questions the answers to which make no conceivable difference, or only the most negligible difference, to the handling of the case. I have often thought that the clinician in charge of the case conference should emulate a professor of law from whom I took a course in equitable remedies, David Bryden. When a law student advanced a stupid argument about the case being discussed, he would respond with a blank stare and the question "And therefore?" This would usuauy elicit some further response from the student (attempting to present the next link in an argumentative chain), but this shoring-up job would in turn be greeted by the same blank stare, the same inquisitorial "And therefore?" I daresay Professor Bryden made the law students nervous; but he also forced them to *think*. I suspect that one who persisted in asking the question "And therefore?" every time somebody made a half-baked contribution to the case conference would wreak havoc, but it might be an educational experience for all concerned.

11. *Ambiguity of professional roles*. When the conference is not confined to one of the three professions in the team, there may arise a sticky problem about roles. For example, in mixed-group conferences I note a tendency to assume that the psychologist's job should be to present the psychometrics and that he is only very gingerly and tentatively to talk about anything else. I think this attitude is ridiculous. I can conduct a diagnostic interview or take a history as well as most psychiatrists, and nonpsychometric data are just as much part of my subject matter as they are of the psychiatrist's. Similarly, if a physician has developed clinical competence in interpreting Rorschachs or MMPI profiles or practicing behavior modification, I listen to what he says without regard to trade-union considerations. By the same token, if I discern that a patient walks with the "schizophrenic float" or exhibits paranoid hyper-vigility or sociopathic insouciance, I feel free to offer this clinical observation in evidence.

12. *Some common fallacies*. Not all of these fallacies are clearly visible in case conferences, and none of them is confined to the case conference, being part of the general collection of sloppy thinking habits with which much American psychiatry is infected. I have given some of them special "catchy" names, admittedly for propaganda purposes but also as an aid to memory.

a. *Barnum effect*. Saying trivial things that are true of practically all

psychiatric patients, or sometimes of practically all human beings—this is the Barnum effect. It is not illuminating to be told that a mental patient has intrapsychic conflicts, ambivalent object relations, sexual inhibitions, or a damaged self-image! (Cf. Meehl, 1956c—reprinted here as Chapter 3; Sundberg, 1955; Tallent, 1958; Forer, 1949; Ulrich, Stachnik, and Stainton, 1963; and Paterson in Blum and Balinsky, 1951, p. 47, and Dunnette, 1957, p. 223.)

b. *Sick-sick fallacy* ("pathological set"). There is a widespread tendency for people in the mental health field to identify their personal ideology of adjustment, health, and social role, and even to some extent their religious and political beliefs and values, with freedom from disease or aberration. Therefore if we find somebody very unlike us in these respects we see him as being sick. The psychiatric establishment officially makes a point of never doing this and then proceeds to do it routinely. Thus, for example, many family psychiatrists have a stereotype of what the healthy family ought to be; and if anybody's family life does not meet these criteria, this is taken as a sign of pathology. Other stereotypes may exist in connection with the "genital character," the person who "fulfills his potential," and so on. Don't let this one pass by, saying that we already know about it! We *do* know about it "officially," but the point is that many people in the mental health field are not very clear about the question in their own thinking. Example: Despite the Kinsey research, some psychiatrists of sexually conservative tastes are likely to overinterpret forms of sexual behavior such as cunnilingus or fellatio as symptomatic of psychopathology, even though the data indicate that mouth-genital contacts have occurred in the *majority* of members of Kinsey's "sophisticated" classes. In my opinion it is almost impossible to say anything clinically significant about a patient on the basis of a history of cunnilingus or fellatio unless one knows a good deal about the motivations. That is to say, it is the motivational basis and not the act which is clinically relevant.

c. "*Me too*" *fallacy* (the unconsidered allegation that "anyone would do that"). This is the opposite of the overpathologizing "sick-sick" fallacy, and one might therefore suppose that clinicians fond of committing the "sick-sick" fallacy would be unlikely to commit the "me too" fallacy. I have no quantitative data on this, but my impression is that the same clinicians have a tendency to commit both. Perhaps the common property is not conservatism or liberalism in diagnosing pathology but

mere sloppy-headedness. The sloppy-headed clinician unconsciously selects, in terms of his personal biases and values, which things he is going to look upon as "terribly sick" and which things he is going to look upon as "perfectly okay" (normal). The example I gave earlier of the nurse who tried to mitigate the diagnostic significance of a patient's visual hallucination by telling us that as a child she had imaginary companions is an example of the "me too" fallacy, although it is compounded with various other errors, such as false analogy and the failure to take developmental stages into account.

I was first forcibly struck with the significance and seductiveness of the "me too" fallacy when I was a graduate student in clinical training. One of my first diagnostic workups was with a girl in late adolescence (a classic Cleckley psychopath: Cleckley, 1964) who was brought in for evaluation on a district court order. She had a considerable history of minor acting out in the form of truancy, impulsive behavior, and running away from home; but the problem which brought her in was that she had "in a fit of pique" hit her foster mother over the head with a lamp base, as a result of which the foster mother sustained a fracture and concussion. One important thing to assess, from the standpoint of the court's inquiry, was the extent to which the patient could exert behavioral control over her impulses. In the 1940's, the patients on our psychiatric service did not have continuous access to their cigarettes but could only smoke at certain times. One of the times when everybody was allowed to come to the nurses' cage to get a cigarette was, let us say, at 3:00 P.M. This particular patient came to the cage around a half hour early and said she wanted her cigarette. The charge nurse told her kindly but firmly that it wasn't quite time yet. The patient insisted that she wanted a cigarette right now and that she didn't want to wait a half hour. The nurse repeated that it wasn't time yet but that she could have a cigarette at 3 P.M. Whereupon the patient began pounding with her fists on the nurse's cage and then flung herself on the floor where she kicked and screamed like a small child having a tantrum. When this episode was discussed in the weekly conference with the junior medical students, the student physician told Dr. Hathaway, the clinical psychologist presiding at the conference, that he didn't see any point in "making a lot out of this tantrum" because, "after all, anybody might act the same way under the circumstances." The dialogue continued thus:

DR. HATHAWAY: "How do you mean 'under the circumstances'?"

MEDICAL STUDENT: "Well, she wanted a cigarette and it's kind of a silly rule."

DR. HATHAWAY: "Let's assume it's a silly rule, but it is a rule which she knows about, and she knows that the tantrum is probably going to deprive her of some privileges on the station. Would you act this way under the circumstances?"

MEDICAL STUDENT: "Sure I would."

DR. HATHAWAY: "Now, think a moment; *would* you, really?"

MEDICAL STUDENT (thoughtful): "Well, perhaps I wouldn't, actually."

And of course he wouldn't. Point: If you find yourself minimizing a recognized sign or symptom of pathology by thinking, "Anybody would do this," think again. *Would* just anybody do it? Behavioristically speaking, what is the actual objective probability of a mentally healthy person behaving just this way? Or, from the introspective point of view, would you *really* do or say what the patient did? Obviously it is *not* the same to say that you might feel an impulse or have a momentary thought similar to that of the patient. The question is, in the case of cognitive distortions, whether you would seriously entertain or believe the thought; or, in the case of overt acting-out conduct, whether you would act out the impulse, having experienced it. You will find that many times, when your initial tendency is to mitigate the symptom's significance in this way, a closer look will convince you that the behavior or belief is actually a serious aberration in reality testing or normal impulse control.

d. *Uncle George's pancakes fallacy*. This is a variant of the "me too" fallacy, once removed; rather than referring to what anybody would do or what you yourself would do, you call to mind a friend or relative who exhibited a sign or symptom similar to that of the patient. For example, a patient does not like to throw away leftover pancakes and he stores them in the attic. A mitigating clinician says, "Why, there is nothing so terrible about that—I remember good ole Uncle George from my childhood, *he* used to store uneaten pancakes in the attic." The proper conclusion from such a personal recollection is, of course, not that the patient is mentally well but that good ole Uncle George—whatever may have been his other delightful qualities—was mentally aberrated. The underlying premise in this kind of fallacious argument seems to be the notion that none of one's personal friends or family could have been a psychiatric case, partly because the individual in question was not hos-

pitalized or officially diagnosed and partly because (whereas other people may have crazy friends and relatives) *I* obviously have never known or been related to such persons in my private life. Once this premise is made explicit, the fallacy is obvious.

e. *Multiple Napoleons fallacy* (the Doctrine of Unreal Realities). This is the mush-headed objection that "Well, it may not be 'real' to us, but it's 'real' to him." (This arises partly from the relativism cultivated by American education or, at a more sophisticated level, from extreme instrumentalism in one's philosophy of science.) It is unnecessary to resolve the deep technical questions of realism and instrumentalism before one can recognize a distinction between reality and delusion as clinical categories. So far as I am aware, even Dewey, Vaihinger, and Heidegger would allow that a man who believes he is Napoleon or has invented a perpetual-motion machine is crazy. If I think the moon is made of green cheese and you think it's a piece of rock, one of us must be wrong. To point out that the aberrated cognitions of a delusional patient "seem real to him" is a complete waste of time. Furthermore, there is some research evidence and considerable clinical experience to suggest that the reality feeling of delusions and hallucinations does differ at least quantitatively, and some investigators allege even qualitatively, from the reality feeling of normal people or from that of the patient regarding familiar nondistorted objects. Thus the statement "It is reality to him," which is philosophically either trivial or false, is also clinically misleading. Nevertheless I have actually heard clinicians in conference invoke this kind of notion on quasi-philosophical grounds, as if to suggest that since nobody knows for certain what reality is, we have no justification for invoking the distinction between the real and the imaginary in assessing a patient.

f. *Crummy criterion fallacy*. It is remarkable that eighteen years after the publication of Cronbach and Meehl's "Construct Validity in Psychological Tests" (1955—reprinted here as Chapter l) and fourteen years after the beautiful methodological development by Campbell and Fiske (1959) and a philosophical treatment by Meehl which has been widely reprinted (1959b—reprinted here as Chapter 5; see also Loevinger, 1957), many clinical psychology trainees (and some full professors) persist in a naive undergraduate view of psychometric validity. (I mention "contemporary" writers—the point about construct validity was made clearly enough by several authors cited in the Cronbach-Meehl paper, and by the great Spearman, whom we unaccountably failed to

mention. It reflects on the shoddy state of psychology that a *graduate* student recently asked me, "Who is this Spearman?") Repeatedly in a clinical case conference one finds psychologists seeing their task as "explaining away" the psychometrics rather than "explaining them" in the sense of genuinely integrating them with the interview, life-history, and ward-behavior material on the patient. It rarely occurs to anyone to feel that he must explain away the intelligence test: the psychiatrist has come to recognize that a successful "bootstraps operation" (Cronbach and Meehl, 1955—see p. 11 above) has been achieved in the measurement of intellect. We do not ordinarily say, "The social worker thought Johnny was dumb, but he has a WISC IQ of 160; isn't it a shame that the test missed again!" But if an MMPI profile indicates strongly that the patient is profoundly depressed or has a schizoid makeup, this psychometric finding is supposed to agree with the global impression of a first-year psychiatric resident, and if it doesn't the psychologist typically adopts a posture of psychometric apology. Now this is silly. Even from the armchair, we start with the fact that an MMPI profile represents the statistical distillation of 550 verbal responses which is considerably in excess of what the clinician has elicited from the patient in most instances, even assuming that the clinician knows how to combine the information he does elicit in an optimal fashion—a proposition at least arguable. Surely there are cases where the psychometrics disagree with the interviewer's clinical impression and yet are at least as likely to be correct as the interviewer, particularly if he is a relatively fresh practitioner in the early stages of his clinical training.

The methodological point is so obvious that it is almost embarrassing to explain it, but I gather it is still necessary. Point: If a psychometric device has been empirically constructed and cross-validated in reliance upon the average statistical correctness of a series of clinical judgments, including judgments by well-trained clinicians as well as ill-trained ones, there is a pretty good probability that the score pattern reflects the patient's personality structure and dynamics better than does the clinical judgment of an individual contributor to the case conference—even if he is a seasoned practitioner, and a fortiori if he is a clinical fledgling. The old-fashioned concept of the "criterion," which applies literally in *forecasting* contexts (such as predicting how much life insurance a person will sell from the insurance salesman key of the SVIB), is not the only appropriate model for the clinical case conference except when we

are explicitly engaged in pragmatic forecasting tasks (e.g., predicting whether the patient will be a continuer or a dropout in outpatient psychotherapy, predicting whether he will respond favorably to Stelazine or EST). It is necessary to be clear about the clinical *task*. Sometimes the clinical task is comparable to the task of the industrial or military psychologist or the educational psychologist trying to select applicants for engineering school who will not flunk out. Most of the time, however, the (alleged) purpose of the clinical case conference is to attain a psychodynamic, nosological, and etiological understanding of the individual patient. I do not enter here into the controversy whether this is an achievable or socially defensible goal, which it may or may not be. The point is that it is the tacitly understood function of much (not all!) of the discussion that goes on in the case conference; given that, it is inappropriate to treat the psychometrics in the same way that we treat them when we have a problem of pure concurrent or predictive validity in the traditional sense.

An MMPI profile is a behavior sample which has been analyzed and summarized in quasi-rigorous fashion on the basis of very extensive clinical experience. This extensive clinical experience has operated first in the construction of the item pool, then in construction and cross-validation of the scales, and then in the development of the various actuarial interpretative cookbook systems. If a patient was diagnosed "reactive depression" by the resident, appears mainly depressed when he is interviewed in the case conference, but has a clearly schizophrenic MMPI supported by some bad schizophrenic F– responses, contamination, and the like on the Rorschach, I cannot imagine why a psychologist would take the simplistic position that his "psychological Wassermann" has failed. If the aim of psychometrics is to help us infer the psychodynamic equivalent of pathology in organic medicine—and that is surely one of its main aims when it is used in a sophisticated way—what the analogy suggests is that there will be, from time to time, discrepancies between what we are prone to infer from the brief interview contact and what Omniscient Jones knows about the psychological innards of the patient.

I don't mean to suggest that we accept the psychometrics as criterion in the old-fashioned sense, which would equally be a mistake. The point is that there *is* no criterion in the traditional sense, and it is preposterous that one still has to explain this to full professors. We do not know the

psychological states and processes from which the various kinds of clinical behavior arise. We *infer* them from a variety of lines of evidence. Our problem is that of the detective (or theory builder!) who is trying to put together different kinds of data to form a more or less coherent picture of unknown latent and historical situations to which he does not have direct operational access. That being so, the task of explaining an apparent discrepancy between the resident's opinion or the impression we get in a case conference and what the MMPI or Rorschach tells us is a much more complicated intellectual job than it seems generally thought to be. As I pointed out in "Some Ruminations on the Validation of Clinical Procedures" (Meehl, 1959b—reprinted here as Chapter 5), giving a Rorschach or an MMPI in order to predict the verbal behavior of the psychiatrist (dynamically or diagnostically) is pointless. It's a waste of the patient's time and the taxpayer's money. If all I want to do is forecast what the psychiatrist will say about the patient's diagnosis or dynamics, it is obvious that the easiest way to do that is to walk down the hall and ask him! A psychometric instrument is not a parlor trick in which, for some strange (union-card?) reason, you keep yourself from having access to easily available information about a patient for the fun of seeing whether you can guess it instead of getting it directly. The psychologist who doesn't understand this point is not even in the ball park of clinical sophistication. To "validate" a test, in any but the crudest sense of initial investigation to determine whether the test has anything going for it at all, a sophisticated thinker realizes that one must use a criterion that is qualitatively and quantitatively superior to what is regularly available in a clinical workup. We validate the Wassermann against the pathologist's and bacteriologist's findings, *not* against the general practitioner's impression after a ten-minute hearing of the presenting complaints. Validation studies that take as the criterion the nosological label or the psychodynamic assessment which one gets on the basis of a couple of interviews are at most always a waste of time. The statements we infer about the patient from psychometrics ought to have attached to them a probability that arises from qualitatively and quantitatively *better* data than we routinely have from the nonpsychometric sources in the ordinary clinical workup. If we don't have that, it is doubtful how much point there is in giving the test in the first place. If a patient has a schizophrenic MMPI and Rorschach but does not appear schizophrenic when interviewed in staff, the *proper*

questions are: "What are some of the things we might have looked for more skillfully to elicit data on the schizoid disposition that the psychometrics indicate are almost certainly present?" "What can be inferred about the psychological defense system of a patient who manages to look like a case of simple depression when he is actually a latent schizophrenic?" "What speculations would we have about discrepancies of this kind?" "What kinds of research might we carry out in order to check these speculations?" "Are there identifiable subclasses of psychometric/ interview discrepancies for which the psychometrics are likely to be correct, and others for which the reverse obtains?" I do not assert that one *never* hears these important metaquestions asked in the case conference; but you can attend a hundred conferences without hearing them raised a dozen times.

g. "*Understanding it makes it normal"* (and, if legal or ethical issues are involved, "acceptable"). This is a psychiatric variant of the ethical notion that understanding behavior makes that behavior ethically permissible or "excusable." I once heard a clinical psychologist say that it was "unimportant" whether a defendant for whom I testified was legally insane, since his homicide was "dynamically understandable" in either case. (The defendant and both counsel, benighted nonpsychologists they, felt it *was* important whether a man is called a murderer and he is put in prison for twenty years or whether he is considered insane and is discharged from the state security hospital after his psychosis lifts.) As for T. Eugene Thompson, the St. Paul lawyer who coldbloodedly murdered his wife to get a million dollars from life insurance, this psychologist argued that "I suppose if I knew enough about T. Eugene Thompson, like the way his wife sometimes talked to him at breakfast, I would understand why he did it." I gather that this psychologist (a Ph.D.!!) believes that if T. Eugene Thompson's wife was sometimes grumpy in the morning, he was entitled to kill her.

h. *Assumption that content and dynamics explain why this person is abnormal*. Of all the methodological errors committed in the name of dynamic psychiatry, this one is probably the most widespread, unquestioned, and seductive. The "reasoning" involved is simple. We find ourselves in possession of two sorts of facts about a person. The first kind of fact, present by virtue of his being a patient, is that he has mental or physical symptoms, or characterological traits, that are pathological in some accepted sense of that term.

This is not the place rigorously to define "pathological," for a beautiful discussion of which see the wise treatment by my colleague William Schofield (1964). For present purposes, it will suffice to say that behavior pathology is roughly defined by some (subjectively) weighted combination of marked statistical deviations from biological and cultural norms, on dimensions and in directions involving (1) subjective distress (anxiety, depression, rage, inadequacy feeling, dissatisfaction, boredom, and the like), (2) medical complaints, symptoms, or concerns, (3) impairment of educational, economic, sexual, or "social" performance, and (4) distorted appreciation of reality, external or internal. It will not usually be the case that any of these aberrations taken alone suffices to define pathology, although there are exceptions involving extreme degrees. For example, no matter how well adjusted socially, economically self-sufficient, and subjectively comfortable a person may be, if he is firmly convinced that he is Napoleon he is pathological ipso facto. It is regrettable, from the standpoint of philosophical cleanness, but the semantic situation must be honestly faced: our conception of psychopathology almost always involves some *mixture* of statistical deviation, "health" or "adjustment" evaluations, and notions of adequate ego function (reality testing and executive competence).

The point is that the individual under study in a clinical case conference comes to be there, unless there has been some sort of mistake (e.g., wrong party in a marriage is the "patient"), because he is psychologically aberrated, i.e., he has psychiatric or medical symptoms, gross social incompetence (delinquency, economic dependency), or extreme deviations in characterological structure. It does not seem useful to define "psychopathology" in solely statistical terms (is absolute pitch, an IQ = 160, or long-sustained sexual performance pathological?). Yet statistical deviations on selected dimensions considered relevant to "health," "social adaptation," "gratification," "effectiveness," and "reality appraisal" seem somehow involved. A down-playing of statistical rarity, in contrast to the work of Schofield cited above, can be found in Fine (1971, pp. 2-6; see also footnote 11 in Livermore, Malmquist, and Meehl, 1968, and citations therein).

The second kind of fact about the person is not true of him by virtue of his being a "patient," but is true of him simply because he is a human being—namely, he has conflicts and frustrations; there are areas of life in which he is less than optimally satisfied, aspects of reality he

tends to distort, and performance domains in which he is less than maximally effective. There is nobody who can honestly and insightfully say that he is always efficient in his work, that he likes everyone he knows ("lie" item on MMPI L scale!), that everybody finds him a fascinating person, that he is idyllically happy in his marriage and his job, that he always finds life interesting rather than boring, that he never gets discouraged or has doubts about "whether it's all worth the trouble," and the like. If you examine the contents of a mental patient's mind, he will, by and large, have pretty much the same things on his mind as the rest of us do. If asked whether there is something that bothers him a lot, he will not emphasize his dissatisfaction with the weather. The seductive fallacy consists in *assuming*, in the absence of a respectable showing of causal connection, that this first set of facts, i.e., the medical, psychological, or social aberrations that define him as a patient, *flows from* the second set, i.e., his conflicts, failures, frustrations, dissatisfactions, and other facts which characterize him as a fallible human being, subject like the rest of us to the human condition. Example: A patient has paranoid delusions that people do not appreciate his merits. He had a father who favored his older brother. One (nonclassical) psychodynamic conclusion is that his present aberrations are mainly attributable to this bit of childhood family dynamics. I do not mean to say that this cannot happen or to deny that sometimes it does. It may be, for all I know, that this inference is true more often than not. By and large, the research literature on retrospective data for persons who have become mentally ill shows only rather weak (and frequently inconsistent) statistical relations between purportedly pathogenic background factors and mental illness (e.g., Schofield and Balian, 1959; Frank, 1965; Gottesman and Shields, 1972). Even those antecedent conditions which do show some association are ambiguous concerning causal interpretation because one does not have any scientific way of determining to what extent the life-history datum—almost always a perception by or of the patient in some interpersonal relation— was itself a reflection of personality aberrations in the "pre-patient" which led siblings, parents, teachers, or peer group to behave differently toward him at an early age. (See, for example, the fascinating study comparing mothers' attitudes toward normal, schizophrenic, and brain-damaged offspring by Klebanoff, 1959.) I do not object to speculating whether a certain event in the patient's past or a certain kind of current

mental conflict *may* have played an important role in producing his present pathological behavior or phenomenology. I merely point out that most of the time these are little more than speculations, whereas the tradition is to take almost any kind of unpleasant fact about the person's concerns or deprivations, present or historical, as *of course* playing an etiological role.

It is worthwhile to distinguish two forms of the mistake in connection with current psychological conflicts or frustrations. The grosser error is to attribute a causal role to an intrapsychic or situational evil when, in the eyes of Omniscient Jones, it has no connection whatever with the presented psychopathology. Thus, for example, a paranoid patient has been out of work for some time due to fluctuations in the economic cycle, and while the development of his paranoid mentation has proceeded quite independently of this unemployment, we assign a causal role to his being out of a job. Sometimes this is done even if the paranoid content itself bears no clear relationship to the alleged situational stressor. But even when it does, the inference remains highly problematic. If I feel put upon by my social environment, I will naturally look around for the most plausible cognitive content in harmony with this feeling; and the fact that I was fired from my job recently is a suitable candidate.

The other form of the mistake is less serious because, philosophically speaking, the alleged factor is really a factor, but its quantitative role is not assigned in a sophisticated manner. These are cases in which a certain factor *does* enter the causal chain eventuating in the pathological symptom which makes the individual classifiable as a mental patient, but it is a factor shared by a very large number—let us say the vast majority—of "normal" persons; and it does not exist in a greater quantitative degree in the patient than it does in the rest of us. The question then arises, why is this particular individual a patient when the rest of us are not? Most often the clarification of such situations lies in the distinction between a genetic or early-acquired disposition and a psychological (environmental) event or condition that appears in the logician's formula as the antecedent term of that disposition. (See Meehl, 1972c—reprinted here as Chapter 11.) Strictly speaking, a *disposition* and *the event that constitutes the realization of its antecedent* count equally as causes. The person can be said to actualize the consequent of the disposition *because* his environment actualizes the antecedent and

because he had the disposition [antecedent → consequent] to begin with, owing to his biological heredity or childhood history. However, when we ask, in a medical or social setting, "What is the matter with this individual?" we do not usually intend to ask, "What is the complete, detailed causal analysis of all the causal chains that converge upon his diagnosably aberrated state as we now see it?" That would be a legitimate question, of course. But it is *not* what we are ordinarily asking when we ask the etiological question "Why?" What we ordinarily have in mind by our etiological "Why?" is "What does this person have, or what befell him, that makes him different from those who have not developed clinical psychopathology?" That means we are looking for the *differentiating* causal agent, the thing which is true of him and not of the others who have remained "healthy." Whether that differentiating agent, picked out of the total causal confluence by our clinical interests, should more properly be the disposition or the realized antecedent term of the disposition depends primarily upon the relative frequencies of the two in the population. If many, perhaps most, persons experience the realization of the antecedent term of the disposition but do not become aberrated because they do not have the disposition to begin with, then the disposition is what is specifically abnormal in this person and should usually be the focus of our clinical and theoretical interest.

The clearest examples of the distinction between the two cases (that is, between a rare disposition whose antecedent is so commonly realized that the antecedent is considered normal and a rarely realized antecedent of a disposition so common that the disposition is called normal) are from medical genetics. In order for a child to develop the PKU syndrome, it is not sufficient that he have a mutated gene at a particular locus, and it is not sufficient that his diet contain phenylalanine. However, the conjunction [mutated gene + dietary phenylalanine] is, given the set of "normal developmental conditions" necessary for the organism to survive at all, jointly necessary and sufficient for PKU (clinical) disease. Why then do we consider this disease hereditary? Obviously, because normal children have considerable phenylalanine in their diet, and the reason they do not develop PKU is that they do not have the mutated gene, i.e., they db not have the disposition. Since the phenylalanine dietary intake is common, PKU is extremely rare, and the reason for its rarity lies in the extreme rarity of the disposition [phenyl-alanine intake → PKU disease], we use the common-language term

"cause" to designate the genetic mutation, i.e., the source of the rare disposition. Comparable examples are diabetes (normal dietary intake of sugar), gout (normal dietary intake of certain nitrogenous foodstuffs), allergies (e.g., normal dietary intake of buckwheat), and the like. And on the other side, the "cause" of lead poisoning or scurvy is taken to be an anomalous dietary intake (excess of lead or deficiency of ascorbic acid), but these are realizations of dispositions that constitute the norm.

There are some circumstances in which, population frequency aside, our choice between the disposition and the realized antecedent as the culprit depends on other contextual parameters, notably therapeutic interest. It may be useful to concentrate our attention upon that which can be changed, irrespective of its rarity. But it is worth noting that in the case of PKU, although we cannot change the child's genes and we can manipulate his diet, any knowledgeable person would unhesitatingly answer the question "Is PKU a genetic disease?" affirmatively. The only basis I can see for this preferential assignment of causality—since a disposition and its actualized antecedent are equally causal in the philosophical sense—is the matter of frequency, i.e., what is the statistically aberrant condition? Expressed in nomic notation, with a genetic (or other constitutional or early-acquired disposition) as '*D*,' the antecedent activation condition of the disposition as '*C*,' and the resulting disease outcome of the combination as '*R*,' the disposition may be written:

$$D = [C \rightarrow R]$$

In our ordinary medical and sociological usage of the term *cause*, with rare exceptions, what we consider is the set of population probabilities $p(D)$, $p(C)$, and $p(R)$. If the relation among these probabilities is

$$p(C) >> p(D) > p(R)$$

we identify the (rare) disposition as the cause; whereas if

$$p(D) >> p(C) > p(R)$$

we instead identify the (rare) actualized antecedent of the disposition as the cause. There is no harm in this selective use of *cause* on the basis of rarity, so long as we are philosophically clear about the situation as thus spelled out. The research tasks in medicine, psychology, criminology, etc., are often profitably put in terms of directing our interest and identification of the cause in this sense of statistical rarity, since one of

the first things we want to know is what it is specifically that is the matter with these individuals, i.e., in what respect do they differ from others who have not fallen ill, have not become delinquent or economically marginal, or whatever.

i. *Hidden decisions*. In practical decision making about patients, it is undesirable to deceive ourselves about those "hidden decisions" that we might challenge were they made explicit, especially that important class of decisions forced upon us by a variety of economic and social factors not presently within our institutional or professional control. An unforced hidden decision is exemplified by the research showing that lower class patients are more likely to receive pills, shock, or supportive therapy than are middle and upper class patients, who are more likely to receive intensive, uncovering, long-term psychotherapy—the latter being, by and large, more congenial to the interests and self-concepts of most practitioners. While this was anecdotally apparent to many of us before it was well documented by Hollingshead and Redlich (1958; see also Myers and Schaffer, 1954), some had supposed that the decision to treat proletarians in a different way hinged almost wholly upon economic considerations. We now know that other factors are also operative, since the social-class correlations persist when economics is substantially eliminated (as at Veterans Administration or other free clinics, graduated-fee community clinics, and the like). These other factors, which should have been obvious to any middle class WASP psychotherapist by introspection, include social-class "cultural compatibility," verbal fluency, conceptual intelligence, the tendency to think psychologically, lesser reliance on somatization (with epinosic gains), less preference for acting-out extrapunitive mechanisms over intropunitive guilt-laden mechanisms, a reality situation that provides some gratification and is modifiable in the nongratifying domains, and the like. Schofield (1964) has described the modal psychotherapist's "ideal patient" as the YAVIS syndrome (young, attractive, verbal, intelligent, and successful).

These YAVIS preferences aside, no practitioner, with or without systematic quantitative research on the sociology of the mental health professions, could be unaware that whether a patient receives a certain kind of treatment—never mind its merits—may hinge negligibly on his objective psychological appropriateness for it, depending instead upon factors of income, geography, available personnel, and the like. It is important in thinking administratively (one may often say also *ethically*)

about the selection of patients for psychotherapy and the assignment of personnel, to face squarely the social fact that even in the affluent society our situation with respect to hours available of professionally skilled time really does present a different situation from that prevailing in other branches of the healing arts. I do not wish to defend the current status of delivery of *non*-mental health care in the United States, which is generally perceived as unsatisfactory. But there are some important quantitative differences between the situation pertaining to psychology and that pertaining to organic disease. Admittedly an indigent patient with a brain tumor may have a significantly lower probability of diagnosis partly because he does not wish to spend money to see a physician about early symptoms, partly because of "social incompetence" traits that show up in caring for one's health (as in all other areas—a social fact that one is not supposed to mention, but is documented by statistical data from prepaid group health care plans). Furthermore, anyone who has gone through (anonymously, not as the "professor" or "doctor" he is) the outpatient department of a charity hospital (something that should be annually required of hospital administrators!) can attest that the underprivileged patient is kept waiting a longer time, is treated with less courtesy and sympathy by paramedical professionals (sometimes scandalously so), is often dealt with rather more high-handedly by the physician, and the like. But despite these conditions, for which there is no excuse, it remains true that the indigent patient, once diagnosed, will not go untreated for his operable brain tumor just because he is poor or because he lives a hundred miles away from the nearest competent neurosurgeon; whereas it is a statistical fact, *not* changeable by some sort of ethical decision or act of will on our part, that the majority of psychiatric patients will not get intensive, long-term psychotherapy (assuming that were the ideal method of treatment for them), money or no money, socially conscious clinic administrator or not, because there are just not enough psychotherapists around.

I have noted in discussion with fellow professionals, and very much in the classroom, that those predictive and prognostic problems that press upon us the clinical-actuarial issue (Meehl, 1954a; Sawyer, 1966) are sometimes rejected with considerable moral indignation, on the plausible-sounding ground that we should not be predicting (fallibly!) who will respond favorably to psychotherapy, since everybody has a

right to it; that we ought to provide it for all comers, even if it happens that their actuarial odds are sometimes rather low for significant improvement. Unfortunately for the clientele but fortunately for the argument, we need not debate the merits of that ethical position—with which I personally have considerable sympathy—because it is a literal, physical impossibility to satisfy this demand, even if all clinical, counseling, and school psychologists, psychiatrists, social workers, clergymen, marriage counselors, and other "mental healers" avoided all teaching and research, and could manage to go without any sleep, recreation, or family life. The situation in psychotherapy is not like the brain tumor, appendicitis, or pernicious anemia situation; it is, regrettably, closer to the situation of a shortage of surgeons or blood plasma in a military field hospital (where overpressed surgeons may literally have to make the decision who shall live and who shall die) or to that of a public health official who runs into a shortage of plague serum during an epidemic of plague. It is not a question of unethically deciding to withhold maximum-intensity psychological treatment from some in favor of others. That decision is already made for us by the sheer logistics of the situation. The point is that we are, willy-nilly, going to withhold intensive psychotherapy from the great majority of persons who come in for some sort of medical or psychological help. Consequently the *character* of our ethical dilemma is fixed. We are not confronted with the problem *whether* to treat some patients intensively and not others. Our present ethical dilemma is whether to assign treatment and nontreatment (or kinds of treatments) in a random fashion or by some selection procedure which improves the average long-term outcome. I cannot think that anyone with a clear head would argue for random assignment (except for research purposes), but I have come across all sorts of strange arguments in this world. In any case, whatever ethical considerations we may raise about the utilization of skilled professional personnel in the foreseeable future, and whatever conclusion we may reach (or agree to disagree on), at least we should keep in mind the fact of hidden decisions.

j. *The spun-glass theory of the mind*. Every great intellectual and social movement seems to carry some "bad" correlates that may not, strictly speaking, follow *logically* from society's acceptance of the "good" components of the movement but that *psychologically* have a tendency to flow therefrom. One undesirable side effect of the mental

hygiene movement and the over-all tradition of dynamic psychiatry has been the development among educated persons (and here I do not refer only to professionals but to many persons who get an undergraduate degree in a variety of majors) of what I call the "spun-glass theory of the mind." This is the doctrine that the human organism, adult or child (particularly the latter), is constituted of such frail material, is of such exquisite psychological delicacy, that rather minor, garden-variety frustrations, deprivations, criticisms, rejections, or failure experiences are likely to play the causative role of major traumas. It is well known among psychotherapists that part of the chronic, free-floating guilt feelings of the educated American woman is her fear that she is not a perfect mother because she is not always 100 percent loving, giving, stimulating, and accepting toward her children. (There is more than a mild suspicion in my mind that some child therapists are ideological "parent haters," drawn to the field by their own parent-surrogate hang-ups.) Some psychotherapists—myself included—actually find it necessary to *undo* the educational and social impact of the mental hygiene movement in women of this sort.

I would do myself a disservice as a clinical practitioner to let these toughminded comments go unqualified. I have a clock on my desk which makes it unnecessary to glance surreptitiously at my wristwatch—one need not hold the spun-glass theory of the mind to notice that checking how close one is to the end of the hour can sometimes have a distinctly adverse effect on patients (particularly schizotypes who, more often than not, react to it as a rejection experience). I offer this minor clinical example to show that I do not here defend a clumsy, insensitive, bull-in-a-china-shop approach to the human psyche. After all, part of the reason people come to psychotherapists is that we offer tact, sensitivity, and empathy beyond that provided by the patient's nurturing environment and by his present family and work group.

Nevertheless, even in one's relations with the patient, it is possible to have a countertherapeutic effect because of subscribing to the spun-glass theory of the mind. The concept of extreme psychic fragility is likely to be truer for the schizotype than for most other kinds of patient, for example. Yet a therapist's *super*-delicacy, flowing from the spun-glass theory of the mind, can boomerang in working with some schizotypes. If, for instance, the therapist is so frightened by the concept "schizophrenia" that he regards it as a kind of psychic cancer, and

therefore tends to react skittishly to some of its major symptoms (e.g., confused thinking, body-image aberrations, reality distortion), he may find himself trying to humor the patient, as "lunatics" are handled in the funny papers, even though all the books and lectures have taught him that this humoring maneuver cannot be successfully carried out. The schizotypic patient, with his hyper-acute perception of others' thoughts and motives—especially when aversive to himself—perceives this therapeutic double-talk as a form of insincerity and feels that the therapist is fooling him while pretending to be honest with him, as, in the patient's view, other people have done in the past. Such an experience confirms the schizotype's deep-seated mistrust, as well as aggravating his cognitive confusions about "what reality is."

The most preposterous example of the spun-glass theory of the mind that has come to my attention illustrates it so beautifully that I can close this portion of my discussion with it. Thirty years ago, when I was an advanced graduate student in Dr. Hathaway's therapy seminar, live-mike interviews were piped in so the staff and students in the class could discuss the therapeutic technique demonstrated. One day we were scheduled to hear an interview by a social worker who (as I had already inferred from other facts) was thoroughly imbued with the spun-glass theory of the mind. The interviewee was a pre-adolescent male with a prostitute mother and a violent, drunken father, living in marginal economic circumstances in a high-delinquency neighborhood, the child having been rejected by his parents, his peer group, and the teachers in his school. His acting-out tendencies and morbid fantasies were such that he was seen on the inpatient child psychiatry service; this session was to be his last interview before discharge, although the social worker planned to continue seeing him with lower density on an outpatient basis. The therapy was considered a success. Shortly before the seminar was scheduled to be held, the social worker informed Dr. Hathaway that she really could not go ahead with the interview as planned, having just learned that the microphone (concealed in a lamp base) was in a different room from the office in which the child was accustomed to being interviewed. She felt that to interview him in this "strange situation" (= different office) might have a traumatic effect and undo the successful achievements of the therapy. This is the spun-glass theory of the mind with a vengeance. Here is this poor little urchin about to be returned to his multiply pathogenic environment, presumably with his

psyche properly refurbished by the interviews so that he will be able to maintain himself in the harsh outside world; yet, despite the "successful" psychotherapy, he is still so fragile that these therapeutic achievements could be liquidated by having an interview in a different office! I submit that the best way to describe that combination of views is that it is just plain silly.

k. *Identifying the softhearted with the softheaded*. While there is surely no logical connection between having a sincere concern for the suffering of the individual patient (roughly, being "softhearted") and a tendency to commit logical or empirical mistakes in diagnosis, prognosis, treatment choice, and the like (roughly, being "softheaded"), one observes clinicians who betray a tendency to conflate the two. Because of my own longtime interest in the clinical-actuarial issue, this is the domain of clinical decision making where the tendency to think and act in terms of the unspoken equation [softhearted = softheaded] has come forcibly to my attention. Given space limitations, its somewhat peripheral relevance, and a firm intention to revise my 1954 monograph (Meehl, 1954a) on the clinical-statistical issue, I shall not reiterate the old arguments—to which, I may say, there have been remarkably few amendments or rebuttals—in the discussion here. But two arguments commonly heard in case conferences bring out the point so beautifully that I cannot resist the impulse to discuss them briefly. One is the old argument that rejects even a strong actuarial prediction concerning the instant patient on the ground that we are concerned not with groups but with this particular individual. Now all predictions about the consequences of clinical action (including inaction, "waiting to see what happens"—often the physician's tactic in accordance with the ancient medical maxim *primum non nocere*) are inherently probabilistic in nature. For one who explicitly recognizes this inherently probabilistic character (even when, as rarely, $p = .99$) of *all* our clinical inferences, the advice to defy our formalized actuarial experience in decision making about the single patient before us amounts to saying that the unformalized inductive inferences of the clinician should be trusted in preference to the formalized probability inferences of a regression equation or an actuarial table. I said in 1954, and have repeated in subsequent publications (Meehl, 1954b, 1956b, 1957—reprinted here as Chapter 4, 1959a, 1960—reprinted here as Chapter 6, 1965c, 1967b—reprinted here as Chapter 9, 1970c, 1972b), that there are individual instances in which

this counteractuarial choice is correct. But I have atso pointed out, and have as yet seen no persuasive rebuttal, that it is very rarely the preferred action and that a policy that permits it frequently is indefensible. Permitting a weak or moderately strong clinical inference to countervail a well-supported actuarial backlog of data on patients resembling the immediate case in a researched set of predictively powerful respects will lead, in the long run, to an increase in erroneous clinical decisions. Some clinicians still do not see that this question is itself one of the questions that is answered, "in the average sense," by the now numerous (over seventy-five) empirical investigations of the clinical-actuarial controversy.

What befalls the softheaded clinician in his admirable desire to be softhearted (i.e., to be most helpful to this particular patient) is that he fears the very real possibility—which the actuarial data themselves express in terms of the error rate—that he will treat the patient nonoptimally through reliance on actuarial experience. I empathize intensely with his existential predicament; I have often felt it acutely myself as a practitioner. But I must insist that he is wrong. In thinking thus, he fails to take two considerations into account. The first is that by departing from the recorded actuarial expectations in reliance upon lower validity informal clinical inferences, he is probably *not* doing the best thing for the immediate case. He thinks (or feels) that he is—but he is probably not. Secondly, should it turn out that by this counteractuarial departure he *has* in fact done the best thing for the particular patient, he will have achieved this individually desirable result by applying a decision policy that (according to the studies) will lead him to mispredict for other patients, who are also individual human beings with presumably as much claim upon his ethical concerns as the one currently before him. In the absence of some showing that we have a kind of superordinate method—whether actuarial or clinical in nature—for discriminating before the fact which are the cases that will be better handled by counteractuarial decisions and which should be left where the table puts them, such a policy is not ethically defensible, regardless of how good it makes us feel.

As to the stock argument that we are not concerned with probabilities, frequencies, or group trends but with the unique individual before us, I do not really know how to add to what I have said, with others before and since, on this vexed issue. There are admittedly some pro-

found unresolved problems, still in dispute among statisticians and logicians, concerning the logical reconstruction of "rational decision" under these circumstances (see, for example, the excellent discussion by Hacking, 1965). But, so far as I am aware, the technical debates among the experts concern the logical reconstruction of the matter, rather than being disputes concerning what a reasonable man would be well advised to do. In teaching our first-year clinical assessment class—where one invariably hears students who offer this "single case" objection to actuarial decision methods in the clinic—I have found it helpful to consider the following hypothetical example (I like this example because it really puts the student on the "existential knife-edge," where he himself is the "patient," and the issue is one of life or death): Suppose I place before you two revolvers. I show you that one of them is loaded with five live shells, having a single empty chamber; the other has five empty chambers and a single live shell. I am, let us say, a sadistic decision-theorist in charge of a concentration camp in which you are an inmate, and I tell you that you are forced to play a single game of Russian roulette with one of these two revolvers. You are not going to have to repeat it. In your ordinary life you are not in the habit of playing Russian roulette. You have never done so before, and you are firmly determined never to do it again. If you avoid blowing your brains out, I promise to release you from the camp. In the other eventuality, we leave the probable outcome to your theology. Which revolver would you choose under these circumstances? Whatever may be the detailed, rigorous, logical reconstruction of your reasoning processes, can you honestly say that you would let me pick the gun or that you would flip a coin to decide between them?

I have asked quite a few persons this question, and I have not yet encountered anybody who alleged that he would just as soon play his single game of Russian roulette with the five-shell weapon. *But why not?* Suppose I am told, by a "softheaded" clinician, "Well, but you are only going to do it once, it is a *unique event*, we are not talking about groups or classes or frequencies—we are talking about whether *you*, Regents' Professor Paul Everett Meehl, that unique human individual, live or die in the next couple of minutes. What do you *care* about probabilities and such, since this choice will never be presented to you again?" I have not found anybody willing to apply such nonactuarial reasoning to the Russian roulette case. Point: We should apply to the unique patient

before us the same kind of rational decision rule that we would insist upon applying if our own life were hanging in the balance.

Despite what I take to be the irrefutability of this two-revolver argument, I can sometimes work myself into the frame of mind of a soft-headed clinician by putting his favorite query, "Do *you* want to be treated as a mere tally mark in an actuarial table?" No, I do not want to be "treated as a *mere* tally mark." But I put it to you, dear reader, that the seductiveness of this appeal lies in a confusion between thinking about my physician's personal concern for my welfare—which I value as highly as anybody else—and trusting him to "bet on the best horse" in my behalf. As a matter of fact, one thing I happen to like about my physician is his tendency (noted appreciatively by other faculty patients of his who are not in the statistics business) to cite statistics when considering whether a certain painful or expensive diagnostic procedure or a certain therapeutic regime is worth trying. I cannot convince myself that it would be a charitable act on my physician's part to think fuzzily about my diagnosis or treatment as a result of his "feeling sympathetic" toward me. Hence I do not think I have a "double standard of morals" that depends upon whether I am considering myself as clinical decision maker or as patient. Whether my physician decides for me, or, as is usually more appropriate—and I would say this also for the psychiatric patient—helps me to decide, I prefer that he act on the principle of Thomas Aquinas that charity is not a state of the emotions but a state of the rationally informed will, i.e., that charity consists of willing the other person's good. On this philosophic basis, it is a *pseudo*charitable act, given the presently available evidence, for a psychiatrist to withhold EST from a patient with classical psychotic depression on the ground that there is something about deliberately inducing a cerebral storm by pushing that button which offends his human sensibility (a feeling I share). By the same token, the psychoanalytic therapist must learn to dissolve resistances rather than timidly playing along with them; an RET practitioner must be able to point out to a proud, educated, intellectualizing patient that he is operating irrationally on a postulate which is unrealistic and self-defeating (tactless though such a confrontation would be in most ordinary human relationships); a behavior modifier must be able to stick to a reinforcement schedule; and the surgeon must not be afraid to shed blood.

It should not require mentioning, but to forestall any possibility of

misunderstanding I shall state explicitly, that all of the foregoing discussion is predicated upon the assumption that a clinical case conference sometimes eventuates in decisions "for" or "about" the patient. Consider the clearly psychotic patient who constitutes a danger to himself or others and whose ego function is so grossly impaired that his relatives (acting through the agency of the state) have placed certain decisions in our hands. One can raise fundamental philosophical questions about such a patient's autonomy in considering the justification of civil commitment (see Livermore, Malmquist, and Meehl, 1968) and if one concludes against current practice, he may have an ethical obligation to refuse to participate in some case conferences, at least in their decision-making aspects. But aside from the involuntary commitment issue, if we do not believe it is a legitimate professional function to decide anything, or even (by advice or by the presentation of relevant information to the patient or his relatives) to help decide anything, then most of the discussion above concerning *how* to decide becomes pointless.

l. *Neglect of overlap*. This one is so trite and has become so much a part of standard elementary instruction in applied statistics that I would have little justification in mentioning it were it not for the almost incredible fact that respectable journals in clinical psychology and psychiatry still persist in publishing articles on the validity of clinical instruments which give no indication that either the author or the journal editor ever heard of the overlap problem. Partly as a result of this "academic" perpetuation of error, case conferences—which usually operate several notches lower in the hierarchy of scholarliness than scientific journals—continue to make the mistake. I suppose the statistics professors are right in their opinion that the primary villainous influence was the unfortunate semantic choice (by whom?) decades ago of the term "significant" in referring to an obtained group difference that cannot plausibly be attributed to random-sampling fluctuations. I am not concerned here with *theoretical* (causal-structure) inferences, commonly made from refutations of the null hypothesis, for a discussion of which see the excellent collection by Morrison and Henkel (1970). The question before us here is the *pragmatic application* of a statistically significant difference, taken for present purposes as being nonproblematic from the statistician's standpoint. The point is that various psychological tests, rating scales, symptom checklists, and the like are unashamedlly proposed for clinical use on the basis of "statistical significance" with

little or no attention paid to the overlap of the clinical populations it is desired to discriminate (assuming that we were to treat the sample statistics not only as establishing a "significant difference" but as infallible estimators). I have repeatedly observed that reminders to faculty and students of the truism that statistical significance does not mean practical importance fail of effect when presented *in abstracto*. At the risk of seeming utterly trivial I shall therefore present a single, simplified numerical example that I hope will carry more pedagogical punch. Suppose I have devised the Midwestern Multiplastic Tennis-Ball Projection Test which I allege to be clinically useful in discriminating schizophrenics from anxiety-neurotics. I set aside the terrible complexities of assessing construct validity for this type of problem, assuming for simplicity that we treat the construct validity as approximately equivalent to a concurrent validity (Cronbach and Meehl, 1955—Chapter 1 above) when the latter has been established on two groups of patients in whose formal diagnoses we are entitled to have much more confidence than we would have on the basis of routine clinical workup (see point f above, "crummy criterion fallacy"). Despite the Fisherian emphasis upon small samples, given our aim to obtain reasonably solid conclusions about the psychometric characteristics of these populations for future use, we would probably be somewhat uncomfortable (if not, we should be!) with sample sizes barely large enough to squeak out a respectable power in refuting the null hypothesis with a *t* test. So let us suppose that we have run the Midwestern Multiplastic Tennis-Ball Projection Test on a carefully diagnosed sample of 100 schizophrenics and 100 anxiety-neurotics. And let us suppose we succeed in achieving a "statistically significant difference" between the two groups at the $p = .01$ level (about par for the course in most journal articles of this sort). To make the computations easy, I shall assume the standard deviations to be equal, and, as indicated above, I shall treat the obtained values as if they were parameters. A little arithmetic applied to these assumptions shows that the ratio of the mean difference $\bar{d}$ to each patient group's standard deviation is approximately .37 which, assuming equal base rates in the clinical population, locates the "hitmax cut" (Meehl, 1973 — Chapter 12 above) midway between the two means, i.e., about .18 sigma units above the mean of the lower frequency distribution and .18 sigma units below the mean of the upper distribution. Entering normal curve tables we find that clinical application of this optimal cutting score

to the dichotomous diagnosis would yield around 57 percent "hits," i.e., a measly 7 percent improvement over what we could achieve by flipping pennies. From my perusal of the current clinical literature I think it not an unfair exaggeration to say that a considerable number—perhaps the majority—of all psychometric differentiators urged upon us for clinical use are close to worthless. A scientific cost accounting of their role in the decision-making process would usually not justify the expense to the patient (or the taxpayer) in the use of skilled clinical time required to administer and score the instrument and to present it in evidence at the case conference.

The conclusion is obvious. We ought to stop doing this sort of silly business, and we should constantly reiterate this elementary point when we note that it has been forgotten by clinicians in the case conference. Also it would be salutary—and would cut down on the garbage found in clinical periodicals—if editors *insisted* that several standard overlap measures be included in every manuscript submitted for publication in which a clinical instrument is purportedly validated or seriously proposed as a device worthy of further exploration. These might be Tilton's overlap, statements of percentages of valid positives attainable by cutting at certain standard percentiles or sigma points on the other distribution (e.g., the median, the 75th percentile, the 90th percentile, the 99th percentile), and, for most clinical problems worth arguing about, an indication of how much employing the hitmax cut on the proposed instrument would be better than "playing the base rate" (Meehl and Rosen, 1955—Chapter 2 above) for various base-rate values.

m. *Ad hoc fallacy*. On this I shall say little at this point because my constructive suggestions for improving the quality of clinical case conferences in Part II below are devoted heavily to this problem. Like the preceding statistical mistake, the ad hoc fallacy is one that everybody "officially" knows about and recognizes as a source of error, but we find it so tempting that we frequently commit it anyway. The ingenuity of the human mind in "explaining" things, the looseness of the theoretical network available to us in the present stage of clinical psychology, and the absence of a quasi-definitive criterion (comparable to the pathologist's report in internal medicine) of what the truth about the patient really is, all combine to make it easy for us to cook up plausible-sounding explanations, after the available relevant evidence is in, of why the patient is the way he is. The only solution to this problem that is

likely to be successful, because it will go beyond mere exhortation and provide quasi-objective differential reinforcement to the verbal behavior of the clinical conferees, is some method that introduces a *predictive* (epistemologically speaking, hence including *postdictive*) element that is now largely lacking. The possibilities that occur to me as reasonably toughminded, not unduly artificial in the pragmatic clinical context, feasible in terms of time and money, and sufficiently enjoyable so that staff can be induced to bear their share of the increased burden, are developed in Part II below.

n. *"Doing it the hard way."* By this I mean employing some clinical instrument or procedure, such as a time-consuming projective test, to ascertain something that documents in the patient's social record or an informant could tell one in a few minutes. I have witnessed tedious and tenuous discussions aimed at making inferences concerning, say, why the patient is an academic underachiever, when nobody had taken the trouble to get in touch with the school and find out how the staff viewed the disparity between his measured intelligence and his academic performance, how the peer group accepted him, what temporal trends showed up in his cumulative record (e.g., teacher ratings), whether he ever was seen by the school counselor, and so on. There are some types of cases in which such failure to look at the record may be especially misleading, such as the clever and ingratiating psychopath who can sometimes fool even a moderately experienced clinician and can completely bamboozle a beginner. Clinicians prone to the [softhearted = softheaded] equation described above, reason, in effect, "Why, this friendly, tousle-headed thirty-five-year-old lad is very cooperative and forms a good relationship with me; I am sure he couldn't have been sticking switchblades into old ladies." In the differential diagnosis between an "unlucky" normal, an acting-out neurotic, a hard-core psychopath, and a solid-gold professional con man, the Rorschach, TAT, and MMPI (or, for that matter, even a short Mental Status interview) may be less illuminating than the school record, a social agency's file, or the police blotter. (See, in this connection, Meehl, 1970a, pp. 10-13.)

In considering psychometrics on their validity, we should try to think clearly about the *role* of our tests in the particular clinical situation. For what purpose are the tests being given? (Of course in thinking about this question, a psychologist who is not clear about the distinctions between content, concurrent, predictive, and construct validity is not up

to the task's demands.) You have to make up your mind *why* you are bothering to give an intelligence test or an MMPI or a TAT. I cannot myself imagine doing so for the purpose of postdicting delinquency, social withdrawal, economic dependency, overdrinking, and the like; but many clinicians seem to view that pointless guessing game as their psychometric task. Just as treating a personality test as a means of predicting some other professional's impressionistic opinion from non-psychometric data is "doing it the hard way," so postdicting a relatively objective fact about the patient's life history is a wasteful exercise in psychometric muscle flexing.

o. *Social scientist's anti-biology bias*. Associated with the spun-glass theory of the mind (as one of the undesirable side effects of the mental hygiene and dynamic psychiatry movements in this country) is a deep, pervasive, and recalcitrant prejudice among psychologists, sociologists, and psychiatrists against biological factors in abnormality. This bias often correlates with a diffuse and fact-blind rejection of biologically oriented treatment procedures. Thus many clinical psychologists are anti-drug, anti-genetic, and anti-EST in their attitudes. Articles and books on psychopathology have been written by eminent and brilliant men (e.g., Thomas Szasz) which not only fail to *refute* the considerable (and rapidly growing) data on genetic determiners of human and animal behavior, including the major psychoses, but—as in the case of Dr. Szasz—do not so much as *mention* in a footnote the existence of such data (see, for example, Erlenmeyer-Kimling, 1972; Gottesman and Shields, 1972; Heston, 1972; Manosevitz, Lindzey, and Thiessen, 1969; and Rosenthal, 1970). One wonders, in reading his writings, whether he is literally unaware of the research on the genetics of schizophrenia; or, if he is aware of it, why he considers it acceptable scholarship to leave the nonprofessional reader in complete darkness about the fact that a scientific controversy exists. For many psychotherapists, everything that is wrong with anybody is attributable either to having a battle-ax mother, being raised on the wrong side of the tracks, or having married the wrong mate. It is dangerous to be the parent or spouse of a mentally ill person because you will almost certainly get blamed for it, even if he was patently abnormal before you met him and his family tree abounds with food faddists, recluses, perpetual-motion inventors, suicides, and residents of mental hospitals. Part of this attitude springs from the two related ideas that if it were the case that

genes had something to do with aberrated behavior, then (1) psychotherapy could not "work," and (2) the psychodynamics we think we understand about mental patients would have to be abandoned. For what I hope is a clear refutation of that undergraduate mistake, see Meehl, 1972c—Chapter 11 above. There simply isn't any contradiction, or even any "friction," between saying in a case conference, "This patient is a schizotype, the specific etiology of which I hypothesize is a dominant gene that produces a specific kind of integrative neural deficit (see Meehl, 1962—Chapter 7 above)" and saying, "This patient's paranoid delusions are restitutional symptoms, forms of miscarried repair the dynamic meaning of which is the patient's effort to reinvest cathexis in social objects." If a clinician thinks that these two statements are incompatible, it merely shows that he is a muddleheaded thinker and needs to take an undergraduate course in genetics plus, perhaps, a little philosophy of science to get clear about dispositions and actualization of their antecedents. Reading Freud will help too.

p. *Double standard of evidential morals*. One common way in which the anti-biological prejudices of the preceding subsection are maintained against contrary evidence is by shifting the standards of evidential rigor depending upon whose ox is being gored. Having been drawn into psychology as a teen-ager by my reading of Menninger, Adler, and Freud, and preferring psychoanalytic therapy (when the patient is appropriate) because it is more theoretically interesting and gives me what I believe to be a deeper causal understanding of the individual, I cannot perceive myself as being a hardnosed, super-rigorous, compulsively operational type of psychologist—although I am aware that the impact of some of my writings on the special problem of prediction has been that other psychologists often view me in this stereotyped way. As mentioned in the introductory section, I have found myself in a strange position vis-à-vis my colleagues: the typical (non-Minnesota) cliniker perceives me as excessively critical and objective, whereas my local psychonomic brethren find it odd that I should be seriously interested in the interpretation of dreams. This is not the place to develop that paradox at length, but in discussing the double standard of evidential morals I must say something about it. I think that one big error committed by psychologists who insist upon sorting other psychologists into boxes like "humanistic" and "scientific" or "dynamic" and "behaviorist" is the failure to distinguish between two sorts of statements.

The first sort of statement is the kind that you might be willing to bet money on, act upon in your personal affairs, rely upon in making decisions concerning a patient—questions on where you place your bets when forced, even though you may be acutely conscious of the fact that you cannot develop the evidence for your choice (when on the existential knife-edge) in a rigorous fashion. The writings on personalistic probability exemplify this (Savage, 1954; Hacking, 1965; Levi, 1967; Raiffa, 1968). There is a difference—but not an inconsistency—between saying, "Lacking coercive evidence, I am prepared, until further notice, to bet that Gallumpher will place in the third," and saying, "It can be shown by rigorous mathematical analysis that the prediction of Gallumpher's placing in the third is the best decision." Consider, for example, psychoanalytic theory. I classify myself as a "60 percent Freudian." I consider that the two men who have contributed most to our understanding of behavior in the first half of the twentieth century are Sigmund Freud and B. Frederic Skinner. I find it a little hard to imagine a conversation between these two geniuses, although I would love to have heard one. But the point is that I can decide, on the existential knife-edge—*required* by the pragmatic context to make decisions willy-nilly—to play it Freudian or Skinnerian, without supposing I can make a rigorous scientific case that my decision is the right one. There is a distinction between what we believe (on the best evidence available, and given the social fact that we *must* decide) and what we would think as pure scientists, which might very well cause us to abstain from any decision until more and better evidence becomes available.

I have no objection if professionals choose to be extremely rigorous about their standards of evidence, but they should recognize that if they adopt that policy, many of the assertions made in a case conference ought not to be uttered because they cannot meet such a tough standard. Neither do I have any objection to freewheeling speculation; I am quite willing to engage in it myself (e.g., I have published some highly speculative views concerning the nature of schizophrenia: Meehl, 1962—Chapter 7 above, 1964, 1972c—Chapter 11 above). You can play it tight, or you can play it loose. What I find objectionable in staff conferences is a tendency to shift the criterion of tightness so that the evidence offered is upgraded or downgraded in the service of polemical interests. Example: A psychologist tells me that he is perfectly confident that

psychotherapy benefits psychotic depressions (a question open on available data), his reason being that his personal experience shows this. But this same psychologist tells me that he has never seen a single patient helped by shock therapy. (Such a statement, that he has never seen a *single patient* helped by shock therapy, can only be attributed to some sort of perceptual or memory defect on his part.) When challenged with the published evidence indicating that shock is a near specific for classical depression, he says that those experiments are not perfect, and further adds, "You can prove anything by experiments." (Believe it or not, these are quotations!) I confess I am at a loss to know how I can profitably pursue a conversation conducted on these ground rules. He is willing (1) to rely upon *his* casual impressions that psychotherapy helps patients, (2) to deny *my* casual impression that shock treatment helps patients, but (3) to reject the controlled research on the subject of electroshock—which meets considerably tighter standards evidentially than *either* his clinical impressions or mine—on the grounds that it is not perfectly trustworthy. It is not intellectually honest or, I would argue, clinically responsible thus to vary your tightness-looseness parameter when evaluating conflicting evidence on the same issue.

I am well aware of a respectable counterargument to these construct-validity considerations, the substance of which is the following: Whatever may be the philosophical or mathematical reconstruction for the idea of construct validity (and the rebuttal is sometimes offered by psychologists who are sophisticated about construct validity as a *theoretical* metanotion), in the pragmatic context whatever we say in the case conference must ultimately come down to some practical decision of a predictive nature. It can even be argued that postdictive, content, and concurrent validity interests—and, a fortiori, construct-validity interests—are defensible in this setting only in reliance upon some relation they have to predictive validity, because the aim of the conference is to decide what to *do* for the patient; this "do" of course includes proposing treatment alternatives to him, making prognostic statements to a referring social institution (court, school), advising the family about the odds on a regime requiring major financial outlay, and the like. In substance, the argument is that whatever the theoretical merits of other kinds of validity, or their technological value over the long run (e.g., improving psychometric instruments through better insight about the construct), in the context of clinical case conferences the *only* kind of

validity that counts is predictive validity. There is much to be said for this line of thought, and no reader familiar with my writings on the actuarial prediction problem would expect me to be unsympathetic to it. And I want to reiterate that there are numerous specific decision-making tasks that do have this pure predictive validity form. Example: A court puts to the professional staff a list of specific forecasting questions, for example, "If the defendant stands trial, will he be able to function well enough cognitively so that his counsel can provide him with an adequate defense?" "This hitherto law-abiding person committed an act of violence under unusual circumstances; if, following your presentence investigation, the court releases him on probation, is he likely to commit acts dangerous to himself or others?" The test of any construct's value in such situations is obviously its predictive power.

Nevertheless, I cannot accept the anti-construct-validity argument when presented in its extreme (hyperoperational) form. My first reservation arises from the social fact that decision making on behalf of the patient or a social institution is not typically the sole function of a clinical case conference. I think it would be generally agreed that the conference is also intended to serve an educational function for the faculty and students attending it.. We are supposedly trying to improve our decision-making skills as helpers and societal advisers, and to clarify our thoughts as teachers and researchers.

In that connection, the display—especially by prestigeful faculty figures—of inefficient decisional procedures must be viewed as countereducational as well as countertherapeutic for the patient. It is not, therefore, even a partial excuse for committing some of the methodological errors I am criticizing to say, "Well, Meehl, you are talking as though the only reason we meet in a clinical case conference is to make decisions about the patient. But we also meet for educational purposes." To the extent that the content of the discussants' contributions is predictive content, fallacies and nonoptimalities in that content, when allowed to go unchallenged or, worse, positively reinforced by group approval, presumably have the effect of indoctrinating our student clinicians with undesirable decision-making habits of mind. Hence the same features that make inefficient decision-making procedures undesirable from the standpoint of helping the individual patient make them undesirable as an educational practice.

The main point I wish to make concerning the educational functions of the conference is that while clinical comments advocating inefficient predictive methods cannot be justified on educational grounds, we are endeavoring to teach the students (and one another) several things in addition to how best to reach concrete clinical decisions about patients for treatment and social forecasting purposes. Admittedly the items in this list of nonpredictive pedagogical aims will differ somewhat from one teacher-professional to another, and I have no wish to impose my hierarchy of personal preferences upon others. I shall merely mention some of the main items that would surely be found in *some* competent persons' lists, without claiming completeness or attempting to argue the merits of the items fully. First, I take it that psychiatrists and clinical psychologists are typically interested in understanding the human person, despite the fact that this understanding does not always lead in any straightforward way to a specific practical decision concerning treatment. I know that this is true for me, and it seems pretty clearly true for many of my colleagues and students. Psychological curiosity is unquestionably among the motives inducing some able minds to enter the profession, and the gratification of *n Cognizance* is for many professionals among the important rewards that keep them going in the face of what is often a somewhat discouraging level of satisfaction of our *n Nurturance*. While some clinicians come fairly close to being pure behavioral engineers, others are more like psychological physicists, the vast majority of us being somewhere in between, characterized by a mixture—sometimes leading to uncomfortable role conflicts—of the wish to *heal*, the wish to *control*, and the wish to *understand*.

I have heard it argued, by extreme representatives of the "tough-minded" end of the tough-tender continuum, that even from the purely theoretical standpoint (setting aside practical relevance in treating the immediate case) this aim to understand cannot be distinguished from the predictive one, since "the purpose of scientific theories is to predict and control." Aside from an element of dogmatism displayed in imposing such a pure instrumentalist view of theoretical science, with which it is possible for a rational man to disagree philosophically, I would emphasize that *some* pragmatically useless inferences may serve epistemologically as corroborators and refuters of nomothetic psychological theories (or their explanatory application to the idiographic material). Such "useless" inferences, when sound, can contribute to the

satisfaction of psychological curiosity without contributing to our role as helpers of patients and social forecasters.

Several kinds of concurrent and postdictive validity illustrate this point. I may, for instance, formulate a construction about the patient's personality by integrating, in the course of the conference discussion, a couple of subtle signs (manifested by the patient when presented in staff conference) with certain aspects of the psychometrics. Relying on this tentative psychodynamic construction, I am led to a probabilistic prediction concerning his ward behavior, which the participant nursing staff then confirms. Assuming that I have not committed any of the methodological errors herein discussed, and that the base rate of my ward-behavior "prediction" (actually postdictive or concurrent validation) is low enough so that its correctness—given the small evidential "prior" p in Bayes' Formula—counts as a strong corroborator; then I have probably learned (and taught) something about this patient's mind and, indirectly, about the verisimilitude of the nomothetic network mediating my inference. But the specificity of treatment in our field is not such that corroborating (in a moderate degree) a particular inference (e.g., this patient has rigid reaction formations against his oral-dependent impulses) must lead directly to a concrete prescription for treatment. The same is, of course, often true for construct-validity-mediated inferences susceptible of confirmation by the patient's psychotherapist.

Again, consider a postdiction which would be, I suppose, largely useless for our helping aim. Suppose I am interested in the theory of depression and entertain the speculation, based partly upon my clinical experience and partly upon quantitative research, that there are at least four, and possibly as many as seven, kinds of depression. Deciding among these for the immediate case *may* have treatment implications; e.g., neurotic depressions and depressions secondary to schizoid anhedonia do not react favorably to EST. But among some of my other speculative depression types, I am not aware of any therapeutic indications. Thus, for example, I believe there is such a state as "rage-depression," and that it even has characteristic somatic complaint aspects not found as frequently in the other varieties, such as the patient's presenting complaint that his head feels as if it had a pressure on it or in it, or as if it were about to explode. These patients also, I believe, are more likely to manifest bruxism. I would contrast this syndrome with

object-loss depression, and would distinguish both from the very common reactive depressions attributable (as Skinner pointed out in 1938) to a prolonged extinction schedule. I speculate that childhood (even adolescent?) object loss predisposes genetically vulnerable persons to subsequent object-loss depression, and the reason it does not show up consistently but only as a statistically significant trend in retrospective studies of depression-prone individuals is that it characterizes only this subgroup (Malmquist, 1970; Beck, Sethi, and Tuthill, 1963; Beck, 1967, Chapter 14). I am not concerned here with arguing the merits of these speculations. The point is that on the basis of the evidence presented in conference, I might be interested in a (quite useless!) postdiction of childhood object loss, whereas in another depressed patient, I might be moved by the way the patient describes his head as feeling as if it were about to explode, together with some violent Rorschach content and some "aggressive" MMPI signs, to inquire whether, according to the patient's wife, he had a tendency to grind his teeth when asleep.

These examples serve also to illustrate the research-stimulus function of the case conference. From the standpoint of research strategy, it may be rational for a research-oriented clinician to find in bits and pieces of concurrent and postdictive validity encouragement to embark upon a research project, although their probabilistic linkage to pragmatically important dispositions of the patient might be too weak to justify reliance upon them in handling the immediate case.

Finally, there is a simple point about construct validity (whether the construct involved is nosological, dynamic, or "historical") that is easy to overlook when our mental set as clinicians emphasizes the importance of predictive statements. A narrowly operational view of the relations between behavioral dispositions (phenotypic, with a minimum of theoretical construction) demands that we have direct evidential support for what would turn out to be an unmanageably huge collection of pair-wise dispositional statistical linkages. If one were to list, in a huge catalog, all of the first-order descriptive traits, signs, symptoms, psychometric patterns, and life-history facts dealt with in psychiatry, it is hardly conceivable that such a list would contain fewer than several hundred elements. Even if we were to prune the list mercilessly—eliminating all elements having (1) marginal reliability, (2) base rates very close to zero or one, or (3) too highly correlated with others having nearly identical "content," and then finally (4) throwing out anything

that we had little or no clinical or research basis for believing was appreciably correlated with anything else we cared about—I find it hard to suppose that such a list would contain fewer than, say, 100 variables. First-order predictions among all these pair wise, if based upon directly researched empirical linkages, would therefore require investigation of 10,000 correlations. But suppose that one investigator finds that bruxism, complaint of exploding headache, and certain MMPI and Rorschach signs cluster as a syndrome which, while "loose," is good enough to provide construct validity for the dynamic nosological entity "rage-depression." And suppose that another investigator, also interested in rage-depression but not familiar with these indicators, reports that patients he and a colleague independently classified as rage-depression (from Mental Status plus psychotherapy evidence plus precipitating situation) respond especially well to a particular antidepressant drug but do badly on Dexamyl. Then, pending the monster study of 10,000 pairwise correlations between everything and everything, clinicians who read these two articles can begin prescribing that specific antidepressant for patients showing the syndrome of bruxism, aggressive psychometrics, and exploding headache.

The same line of reasoning applies to the teaching of diagnostic, dynamic, and etiological factors. Presumably one justification for having case conferences instead of just sending all of the residents and psychology trainees to the library is our belief that certain things can be best taught with dramatic punch in the real-life clinical situation. I do not know whether that generally accepted pedagogic principle has been quantitatively researched in medicine, but the psychiatric and clinical psychology conference has accepted the tradition from other branches of medicine, and I am willing here to presuppose it. You can "tell" a resident or psychology trainee that many schizophrenic patients are baffling and frustrating to the therapist, and elicit adverse countertransference reactions not because the therapist has been technically mishandling the case—although he may now begin to do so!—but because the schizotype is prone to "testing" operations on persons he would like to trust but dare not. You can also state in a lecture that some schizophrenic patients have a special way of walking (I will not try to describe it verbally here) which I refer to as the "schizophrenic float." A fledgeling therapist, mistreating a pseudoneurotic schizophrenic as a "good healthy neurotic," comes into the conference hurt and puz-

zled by the patient's ambivalent testing operations. The schizophrenic float is called to the therapist's attention by his conference neighbor (who spots it as the patient walks in), and the student therapist has a chance to observe it as the patient leaves the conference room. This resident or psychology trainee will have formed a vivid connection in his clinical thinking that it is likely he will never forget. However, such a linkage need not be formed on the same patient, although it's better that way. If the senior staff succeed in convincing the resident in this week's conference that the reason for his countertransference troubles lay in the patient's being a pseudoneurotic schizophrenic, and next week he sees some other student's patient showing the schizophrenic float as he walks into the room, that pair of experiences will perhaps do almost as well.

13. *Antinosological bias*. It is common knowledge that American psychiatry and clinical psychology, the former under psychodynamic influence and the latter under both psychodynamic and learning theory influence, have an animus against formal "diagnosis." The status of formal nosological diagnosis in American theory and practice warrants detailed treatment, and I am preparing such a discussion of theory and research literature for presentation elsewhere. I shall therefore confine myself here to a mere listing of some of the current clichés, with brief critical comments upon each but without attempting an adequate exposition of the argument or—when decent empirical data exist—detailed survey of the research findings. There are, of course, good reasons for being skeptical about diagnostic rubrics, and even more skeptical about their current application in a psychiatric tradition that deemphasizes training in diagnostic skills. But it is regrettable to find that the majority of beginning graduate students in clinical psychology "know" that "mere diagnostic labels" have no reliability or validity, no theoretical significance, no prognostic importance, and no relevance to treatment choice. They "know" these things because they were told them dogmatically in undergraduate abnormal psychology classes. They typically react with amazement, disbelief, and resentment to find a psychologist who bluntly challenges these ideas. If you want to be a diagnostic nihilist, you should be one in an intellectually responsible way, for scientific reasons rather than from bobbysoxer antidiagnostic propaganda. On the current scene, antidiagnostic prejudices of the familiar kinds (four of which I consider here) have recently been bolstered by a new ideological factor,

to wit, the tendency of many students to *politicize everything*. A professor can (perhaps) discuss the helium nucleus or the sun's temperature without finding himself shortly involved in a debate on women's liberation, police brutality, Indochina, "establishment" bourgeois values, or the black ghetto. But psychiatric diagnosis is one of those topics that are reflexly politicized by many of our students.

Herewith, then, a brief summary of the usual antidiagnostic arguments, and my objections to each:

a. "Formal diagnoses are extremely unreliable." If it were empirically shown that formal diagnoses are extremely unreliable, it would remain an open question whether they are unreliable because (1) the diagnostic constructs do not refer to anything that really exists (i.e., there *is* no typology or taxonomy of behavior aberration that "carves nature at its joints,"), or (2) differential diagnosis of behavior disorders is unusually difficult, or (3) it is not unusually difficult but many clinicians perform it carelessly and uninformedly. One ought not, after all, be astounded to find that American psychiatrists and psychologists, educated in an antinosological tradition in which they have been taught that diagnosis is of no importance (and consequently never exposed to the classic nosological writings in the European tradition), have been presented with professional models of senior staff who do not take diagnosis seriously, and have not been differentially reinforced for good and poor diagnostic behavior, are unable to do it well!

It is not true that formal nosological diagnosis in psychiatry is as unreliable as the usual statements suggest. If we confine ourselves to major diagnostic categories (e.g., schizophrenia versus nonschizophrenia, organic brain syndrome versus functional disorder, and the like), if we require adequate clinical exposure to the patient (why would anyone in his right mind conduct a study of diagnostic rubrics based upon brief outpatient contact?), and if we study well-trained clinicians who take the diagnostic process seriously, then it is not clear that interclinician diagnostic agreement in psychiatry is worse than in other branches of medicine. (A colleague responds with "That's true, but medical diagnoses are completely unreliable also." I am curious what leads this colleague, given his "official" classroom beliefs, to consult a physician when he is ill? Presumably such an enterprise is pointless, and taking your sick child to a pediatrician is wasted time and money. Do any of my readers *really* believe this?) For instance, as to the diagnostic

dichotomy schizophrenia versus nonschizophrenia, one study—based upon a very large *N*—shows the interjudge reliability to equal that of a good individual intelligence test (Schmidt and Fonda, 1956). I do not mean to suggest that the various interjudge reliability studies are consistent, which they are not (see, for example, Rosen, Fox, and Gregory, 1972, Table 3-1, p. 46); nor do I assert that the evidence on this question is adequate at present. I merely point out that the majority of psychologists and psychiatrists in this country persist in reflexly repeating the dogma "Diagnosis is very unreliable" without paying due attention to the diagnostic circumstances and personnel involved in various studies, or telling us how unreliable something has to be before it is "very unreliable." The spectacle of a clinical psychologist spitting on formal psychiatric diagnostic labels on grounds of unreliability, meanwhile asking us to make clinical decisions on the basis of Rorschach interpretations, can only be described as ludicrous." For an excellent survey and sophisticated criticism of the empirical research on diagnostic interjudge reliability, plus some impressive oew data on the subject, see Gottesman and Shields (1972, Chapter 2). I need hardly add that the errors criticized in this paper are presumably a major source of diagnostic unreliability, so that their reduction would yield an improvement (I predict a big improvement) over typical reported coefficients.

b. "We should be interested in understanding the patient rather than labeling him." This muddleheaded comment may be given additional controversial power by describing a taxonomic rubric as a "pigeonhole," whereby a clinician who diagnoses his patients or clients is adjudged guilty of "putting people into pigeonholes"—a manifestly wicked practice, the wickedness being immediately apparent from the very words, so no further argument is required. *Res ipsa loquitur!*

It should not be necessary to explain to sophisticated minds that whether "labeling" in the nosological sense is *part* of "understanding" the patient cannot be decided by fiat, but hinges upon the etiological content of the label. If the nosological label is a completely arbitrary classification corresponding to nothing in nature, then it is admittedly not contributory to our understanding the patient we are trying to help. And of course if that is its status, it is not contributory to anything (even epidemiological statistics) and shouldn't be engaged in. Anyone who uses formal nosological categories responsibly should, in consistency, believe that the rubrics mean something. (He need not, obviously,

believe that they *all* mean something.) A diagnostic label means something about genetics or salient conflicts or schizophrenogenic mothers or social-class factors or unconscious fantasies or preferred mechanisms of defense or aberrated neurochemistry or whatever; and these kinds of entities are aspects—frequently clinically relevant aspects—of an adequate "causal understanding." It is important to see that which *class* of theoretical entities is implied by the nosologic term still remains open after a methodological decision to permit nosological labels is made. To conflate the two questions—"Are there taxonomic entities in psychiatry?" and "Is aberrated behavior sometimes caused by germs, genes, or structural CNS conditions?"—is just dumb, but the conflation is well nigh universal in American clinical thinking. See, in this connection, Meehl, 1972c—chapter 11 above; also footnote 19 (at p. 80) of Livermore, Malmquist, and Meehl, 1968; and footnote 10 (at p. 12) of Meehl, 1970b. The widespread habit of mentioning the "medical model" without having bothered to think through what it is (causally, statistically, and epistemologically) prevents an intellectually responsible consideration of complex taxonomic questions. An "organic" causal factor (e.g., vitamin deficiency) may be taxonomic or not; so also for a genetic causal factor (e.g., PKU mental deficiency is taxonomic, but garden-variety hereditary stupidity is not). On the other hand, a "nonorganic, nongenetic," purely social-learning etiology, while perhaps *usually* non-taxonomic, may sometimes be taxonomic. The schizophrenogenic mother has been so conceived by some. Suppose that Freud had been correct in his (pre-l900) opinion about the respective etiologies of hysteria and the obsessional neurosis. He held, on the basis of his early psychoanalytic treatment of these two groups of patients, and before his shattering discovery that much of his psychoanalytic reconstruction of their early childhood was fantasy, that the specific life-history etiology of hysteria consisted of prepubescent sexual (specifically *genital*-stimulation) experience in which the future patient was passive and in which fear or disgust predominated over pleasure. Whereas he thought that the obsessional neurosis had its specific life-history origin in prepubescent sexual experience in which the subsequent patient played an active (aggressive) role and in which pleasure predominated over the negative affects. Had this specific life-history etiology been corroborated by subsequent investigation, the diagnostic labels "hysteria" and "obsessional neurosis" would have carried a heavy freight of causal

understanding, and would have been truly taxonomic. It makes no difference what *kind* of etiology we focus on (social, genetic, biochemical, or whatever), so long as the label points to it.

The notion that subsuming an individual under a category or rubric somehow prevents us from understanding the causal structure of his situation is one which has been repeatedly criticized but with negligible effect. The methodological level at which such discussions are typically carried on in the American tradition is pathetic in its superficiality. So far as I can discern, most clinicians who talk about the subject in this way have never even asked themselves what they *mean* by saying that "There are no disease entities in functional psychiatry." To make such a negative statement significantly, one ought presumably to have some idea about what would be the case if there *were* "entities" in functional psychiatry. One cannot deal with complicated questions of this sort by a few burblings to the effect that schizophrenia is not the same kind of thing as measles. What kinds of causal structures (and resultant phenotypic correlations and clusterings) may conveniently be labeled as "real entities" is a metaquestion of extraordinary complexity. To think about it in an intellectually responsible way requires philosophical, mathematical, and substantive competence at a level possessed by very few psychiatrists or clinical psychologists. Much of what we have to think clearly about in connection with the nosology-dynamics problem is tied up with the genetic factors problem in psychodynamics (cf. Meehl, 1972c—chapter 11 above).

c. "Formal diagnoses are prognostically worthless." This statement is just plain false as a matter of empirical fact. No one familiar with the published statistics, and for that matter no one who has kept his eyes and ears open around a mental hospital for a while, can deny—unless he has been brainwashed into a rabid antidiagnostic prejudice that paranoid schizophrenia has a very different outlook from a nice clean hypomanic attack in a cyclothymic personality, or that a "reactive depression" (precipitated, say, by failing one's Ph.D. prelims) will run a shorter course (with or without psychotherapy or chemotherapy) than a textbook compulsion neurosis, or that a hard-core Cleckley psychopath (Cleckley, 1964) is likely to continue getting into trouble until he becomes old enough to "simmer down" or "burn out," or that a case of hypochondriasis has a very poor outlook. I find it strange that psychologists urge us to rely upon psychological tests (especially the low-

validity projective methods) for predictive purposes when, so far as the record shows, they do not have as much prognostic power as does formal diagnosis *even when made sloppily as in this country*.

Consider such a life-or-death prognostic problem as suicide risk in patients suffering from psychotic depression. Despite Bayes' Formula, and the arguments advanced by my doctoral student and co-author Albert Rosen in his paper on suicide (Rosen, 1954; see also Meehl and Rosen, 1955—Chapter 2 above), in cases of psychotic depression the suicide risk figure is large enough to take into serious account. The usual estimates are that, before the introduction of EST and the antidepressant psychotropic drugs, roughly *one psychotic depression in six managed to kill himself*. (This figure cannot, of course, be easily calculated from the usual epidemiologic "rate" value.) More recently, follow-up studies of psychotically depressed patients who had made a "clinical recovery" sufficient to be discharged from the hospital found that another 3-5 percent will commit suicide in the ensuing two or three years after discharge. Point: Suicide probability among patients with psychotic depression is approximately equivalent to death risk in playing Russian roulette. If the responsible clinician does not recognize a psychotically depressed patient as such, and (therefore) fails to treat him as having a suicide risk of this magnitude, what he is in effect doing is handing the patient a revolver with one live shell and five empty chambers. Considering the irreversibility of death as an event, and the disutility attached to it in our society's value system, I assume my readers will agree that a Russian roulette probability figure is nothing to treat cavalierly. *Any psychiatrist or psychologist who does not make a thorough effort to ascertain whether his patient has a psychotic depression rather than a "depressive mood"* (the most common single psychiatric symptom, found in a wide variety of disorders), *in order to determine whether the patient requires treatment as a suicide risk of this magnitude, is behaving incompetently and irresponsibly*.

I will add some punch to this statistical argument by relating an anecdote (it comes to me directly from the student clinician to whom it happened). I report it in the form of a dialogue between myself and the student. This student therapist (a "psychiatric assistant") is an extremely bright, highly motivated, and very conscientious Arts College senior with three majors (one of which is psychology) and an HPR = 3.80. I mention these facts as evidence that the student's ignorance

arises *not* from stupidity, lack of curiosity, poor motivation, or irresponsibility. It arises from the antinosological bias (more generally, the antiscientific, anti-intellectual attitudes) of his teachers and supervisors. The exchange goes as follows:

MEEHL: "You look kind of low today."

STUDENT: "Well, I should be—one of my therapy cases blew his brains out over the weekend."

MEEHL: "Oh, I'm sorry to hear that—that is a bad experience for any helper. Do you want to talk about it?"

STUDENT: "Yes. I have been thinking over whether I did wrong, and trying to figure out what happened. I have been his therapist and I thought we were making quite a bit of progress; we had a good relationship. But then he went home on a weekend pass and shot himself."

MEEHL: "Had the patient talked to you about suicide before?"

STUDENT: "Oh, yes, quite a number of times. He had even tried to do it once before, although that was before I began to see him."

MEEHL: "What was the diagnosis?"

STUDENT: "I don't know."

MEEHL: "You mean you didn't read the chart to see what the formal diagnosis was on this man?"

STUDENT: "Well, maybe I read it, but it doesn't come to my mind right now. Do you think diagnosis is all that important?"

MEEHL: "Well, I would be curious to know what it says in the chart."

STUDENT: "I am not sure there is an actual diagnosis in the chart."

MEEHL: "There *has* to be a formal diagnosis in the chart, by the regulations of any hospital or medical clinic, in conformity with the statistical standards of the World Health Organization, for insurance purposes, and so on. Even somebody who doesn't believe in diagnosis and wouldn't bother to put it in a staff note must record a formal diagnosis on the face sheet somewhere. He has to put something that is codeable in terms of the WHO *Manual of the International Statistical Classification of Diseases, Injuries, and Causes of Death*."

STUDENT: "Oh, really? I never knew that."

MEEHL: "Did you see this man when he first came into the hospital?"

STUDENT: "Yes, I saw him within the first week after he was admitted."

MEEHL: "How depressed did he look then?"

STUDENT: "Oh, he was pretty depressed all right. He was very depressed at that time."

MEEHL: "Well, was he psychotically depressed?"

STUDENT: "I don't know how depressed 'psychotically depressed' is. How do you tell a psychotic depression?"

MEEHL: "Hasn't anybody ever given you a list of differential diagnostic signs for psychotic depression?"

STUDENT: "No."

MEEHL: "Tell me some of the ways you thought he was 'very depressed' at the time he came into the hospital."

STUDENT: "Well, he was mute, for one thing."

MEEHL: "*Mute?*"

STUDENT: "Yes, he was mute."

MEEHL: "You mean he was not very talkative, or do you mean that he wouldn't talk at all?"

STUDENT: "I mean he wouldn't talk at all—he was mute, literally mute."

MEEHL: "And you don't know whether that tells you the diagnosis—is that right?"

STUDENT: "No, but I suppose that means he was pretty depressed."

MEEHL: "If he was literally mute, meaning that he wouldn't answer simple questions like what his name is, or where he lives, or what he does for a living, then you have the diagnosis right away. If the man is not a catatonic schizophrenia, and if you know from all the available evidence that he is some kind of depression, you now know that he is a psychotic depression. There is no such thing as a neurotic depression with muteness."

STUDENT: "I guess I didn't know that."

MEEHL: "Why was he sent out on pass?"

STUDENT: "Well, we felt that he had formed a good group relationship and that his depression was lifting considerably."

MEEHL: "Did you say his depression was *lifting*?"

STUDENT: "Yes, I mean he was less depressed than when he came in—although he was still pretty depressed."

MEEHL: "When does a patient with a psychotic depression have the greatest risk of suicide?"

STUDENT: "I don't know."

MEEHL: "Well, what do the textbooks of psychiatry and abnormal psy-

chology *say* about the time of greatest suicide risk for a patient with psychotic depression?"

STUDENT: "I don't know."

MEEHL: "You mean you have never read, or heard in a lecture, or been told by your supervisors, that the time when a psychotically depressed patient is most likely to kill himself is when his depression is 'lifting'?"

STUDENT: "No, I never heard of that."

MEEHL: "Well you have heard of it now. You better read a couple of old books, and maybe next time you will be able to save somebody's life."

The obvious educational question is, how does it happen that this bright, conscientious, well-motivated, social-service-oriented premed psychology major with a 3.80 average *doesn't know the most elementary things about psychotic depression*, such as its diagnostic indicators, its statistical suicide risk, or the time phase in the natural history of the illness which presents the greatest risk of suicide? The answer, brethren, is very simple: Some of those who are "teaching" and "supervising" him either don't know these things themselves or don't think it is important for him to know them. This hapless student is at the educational mercy of a crew that is so unscholarly, antiscientific, "groupy-groupy," and "touchy-feely" that they have almost no concern for facts, statistics, diagnostic assessment, or the work of the intellect generally.

d. "Diagnosis does not help with treatment." This is, of course, not a valid criterion for determining whether formal diagnoses have factual meaning, empirical validity, or interjudge reliability; that it is even thought to be so reflects the shoddy mental habits of our profession. But its conceptual implications aside, how much truth is there in the assertion, given the baselines of accuracy in treatment choice we generally have to live with in clinical psychology? I would be interested to learn that any psychological, test, or any psychodynamic inference, has a treatment selection validity as high as the nosological distinction between the affective psychoses and other disorders with regard to the efficacy of one of the few near-specifics we have in psychiatry, to wit, EST. Even a much less specific treatment indication, the phenothiazines for schizophrenia, has, as I read the record, as good a batting average as psychometrics or psychodynamic inference (see, for example, Meltzoff and Kornreich, 1970; Bergin and Strupp, 1972).

As elsewhere in this paper, I have here the occasion to point out the

problem of a "double standard of methodological morals." If somebody is superskeptical and superscientific and requires reliability coefficients regularly better than .90 before he will use a proposed category or dimension in clinical decision making, then he will have a hard time justifying formal psychiatric diagnosis even when it is made by well-trained diagnosticians. (He will also have to advocate that physicians abandon their pernicious habit of taking blood pressures!) But such a superskeptic ought not, in consistency, waste his or our time in a case conference gassing about the patient's family dynamics or his unconscious mechanisms or his Rorschach or TAT or MMPI—*because none of these, singly or collectively, can measure up to his strict methodological demands either*. The decrying of diagnosis by psychiatrists and psychologists in favor of psychodynamic understanding or psychologist's test interpretation would require a showing that these competing methods of prediction and treatment choice are superior to psychiatric diagnosis when each is being done respectably. So far as I have been able to make out, there is no such showing

Part II: Suggestions for Improvement

The preceding discussion has admittedly been almost wholly destructive criticism, and I confess to having written it partly motivated by a need for catharsis. Being an oral-impatient character with a 99th percentile "theoretical" score on the Allport-Vernon-Lindzey Values, my boredom tolerance is regrettably low. I don't really mind it much when my colleagues or students ignore me or disagree (interestingly) with me—but I become irritated when they bore me. It is annoying to walk across campus to the hospital and find oneself treated to such intellectual delicacies as "The way a person is perceived by his family affects the way he feels about himself—it's a dynamic interaction," or "Schizophrenia is not like mumps or measles." However, having expressed some long-standing irks and, I hope, having scored a few valid points about what is wrong with most case conferences in psychiatry or clinical psychology, I feel an obligation to try to say something constructive. Not that I accept the pollyanna cliché that purely destructive criticism is inadmissible. This has always struck me as a rather stupid position, since it is perfectly possible to see with blinding clarity that something is awry without thereby being clever enough to know how to cure it. I

do not know how to stop religious wars or structural unemployment or racial prejudice or delinquency or divorce or mental illness—but I am tolerably clear that these are undesirable things in need of amelioration. Whether the following proposals for improving the quality of clinical case conferences are sound, about which I have no firm opinion, does not affect the validity of the preceding critical analysis. I invite the reader who does not find himself sympathetic to my proposed solutions to look for alternative solutions of his own.

The first suggestion that comes to the mind of anyone whose training emphasized differential psychology (and I am old-fashioned enough to believe that trait analysis is still important) is an improvement in the intellectual caliber of the participants. Obviously this is not something one can go about accomplishing directly by administrative fiat. We can't pass an ordinance requiring of the cosmos that more people should have super-high IQ's! However, several top schools (Minnesota included) have in recent years opted for a marked reduction in size and goals of their Ph.D. clinical psychology training programs, which has permitted the imposition of tougher "scholarly standards." The social issues involved are vexatious and beyond the scope of this paper.

More difficult to assess quantitatively, and therefore more subject to my personal biases, is the question of value orientation, what "turns people on." In my graduate school days, those of my peers who went over to the University Hospitals to work on the psychiatric ward and with Dr. Hathaway on MMPI development were students having both a strong interest in helping real flesh-and-blood patients (not to mention the fun of wearing a white coat!) *and* intense cognitive hungers. While wanting to be clinicians, they were characterized by "intellectual passion"; they would all rate very high on *n Cognizance*. But most observers of the contemporary psychological scene agree with me that strong cognitive passions (and their reflection in highly scholarly achievement and research visibility) have, alas, a distinct tendency to be negatively associated with a preference for spending many hours per week in service-oriented, face-to-face patient contact. This anecdotal impression (noted by every psychologist I have asked) receives indirect quantitative support in the well-known negative correlations (many in the −.50's and −.60's, some in the −.80's) between "scientific" and "uplift" scores on the SVIB (Strong, 1943, Table 193, p. 716; Campbell, 1971, Table 2-4 on p. 36, Table 3-31 on p. 111); the weak "so-

cial" tendencies of Terman's gifted subjects as children and adults (Terman, 1925, p. 420; Burks, Jensen, and Terman, 1930, pp. 173- 176; Terman and Oden, 1947, pp. 36-37; Terman and Oden, 1959, p. 10; see also Hollingworth, 1926, *passim*); Robert Thorndike's investigations of activity preferences and values of psychologists (Thorndike, 1954, 1955; see also Clark, 1957, pp. 85, 90-95, 112, 224-225; related are Shaffer, 1953, and Campbell, 1965). Highly creative professionals have been shown in several studies to be less "socially oriented" than uncreative controls (see, for example, Dellas and Gaier, 1970, and references cited therein). But this negative correlation between "social" and "cognitive" passions is very far from being perfect. Hence we can select, *if we have a rather small N of trainees in a program*, applicants falling in the (++) cell of a cognizance-nurturance fourfold table. However, when the N becomes very large, when the particular psychology department has a reputation for an "applied emphasis," and when the criteria of selection become somewhat less scientifically or intellectually oriented, then one finds an increasing number of trainees in the program who are really not "turned on" by the life of the intellect. These students, admirable as human beings and doubtless well-motivated healers, find themselves somewhat bored, and in some cases actively irked, by abstract ideas.

As I said above, I am somewhat old-fashioned in these matters. I believe there is no substitute for brains. I do not believe the difference between an IQ of 135—perfectly adequate to get a respectable Ph.D. degree in clinical psychology at a state university—and an IQ of 185 is an unimportant difference between two human beings (cf. McNemar, 1964). Nor do I believe a person, even if basically bright, can be intellectually *exciting* unless he is intellectually *excited*. It astonishes me that so many persons enter academic life despite having what, to all appearances, is a rather feeble capacity for becoming excited about ideas. This aspect of the case conference problem—the fact that many of its participants are not first-class intellects in either ability or values— is obviously not curable by any modification in format.

However, without being unkindly elitist, we might try to convey (gently but firmly) the message that if you don't have anything worthwhile to say, you should probably shut up. The current practice, based upon a kind of diffuse "T-group" orientation to case conferences, seems to assume that everybody should get into the act regardless of how bright

he is or what he knows, either clinically or theoretically. I view this attitude as preposterous on the face of it. The plain fact is that what most people have to say about anything complicated is not worth listening to. Or, as my medical colleague Dr. Howard Horns put it in a lovely metaphysical witticism, "Most people's thoughts are worth their weight in gold." If it is argued that you can't prevent people who have nothing significant to contribute from talking without being cruel or discourteous, I submit that this is empirically false. I point to case conferences in other specialties like neurology and internal medicine, where, so far as I have observed, there is no social discourtesy or cruelty manifested by those in charge; but the general atmosphere is nevertheless one which says, in effect, "Unless you know what you are talking about and have reason to think that you are saying something really educational for the rest of us or beneficial to the patient, you would be well advised to remain silent. Mere yakking for yakking's sake is not valued in this club." I have rarely had to listen to trivia, confused mentation, plain ignorance, or irrelevancies when I have attended case conferences in internal medicine or neurology, or the clinicopathological conference on the medical service. If an atmosphere of decent intellectual scholarly standards can be created and maintained on those services, I cannot think it is impossible to approximate the same thing in clinical psychology and psychiatry.

Mention of the clinicopathological conference in medicine brings me to my tentative and sketchy suggestions for improving the *format* of the case conference, suggestions largely although not entirely independent of the two preceding (unchangeable?) factors. One of the main reasons why so much hot air is emitted and reinforced in the psychiatric conference has almost nothing to do with the intellectual competence of the participants, namely, the sad fact that nobody can be proved wrong in what he says because there are no even quasi-objective external criteria. As is well known, one of the great contributions of Dr. Richard Clarke Cabot in dreaming up the clinicopathological conference—reports on the conferences from the Massachusetts General Hospital still appear regularly and would be highly educational reading for clinical psychologists, to whom I recommend the collections (Castleman and Burke, 1964; Castleman and Dudley, 1960; Castleman and McNeill, 1967; Castleman and Richardson, 1969)—is that everybody is put on the spot. For instance, a distinguished, world-famous visiting professor of

medicine might be asked, on the basis of the clinical material presented, to set up the differential diagnosis, to argue the pros and cons, to ask for additional data that may not have been presented in the first go-around, and finally to stick his neck out and make a guess about what the pathologist found postmortem. While pathology is not, strictly speaking, a definitive operational criterion in the logician's sense (as anyone can easily discover by attending a clinicopathological conference in, say, pediatrics and listening to the pathologists debate whether the blood-cell slides are or are not early leukemia of a certain kind), still, for many diseases, the pathological findings can be taken as quasi-criteria. No matter what kind of psychiatric and neurological symptoms a patient shows clinically, if he has a negative blood and spinal fluid Wassermann, if his cerebral cortex does not show characteristic paretic changes, and if his brain tissue is completely free of *Treponema pallidum*, then he does not have paresis. If all the neurologists had agreed "clinically" that he was paretic, the interesting questions in the conference then become "What did he have instead?" and "How were the clinicians led so badly astray?" Point: A clinicopathological conference in neurology or medicine is an educational experience for students and staff largely because there is a *right answer*. And one desirable fringe benefit of the existence of this quasi-criterial "right answer" is the non- reinforcement of foolish conversation. If you say something grossly stupid, you are almost certain to be found out when the pathologist enters the fracas at the end of the conference.

A diagnostic entity in organic medicine is quasi-defined by the conjunction etiology-cum-pathology. If there were microscopically and chemically indistinguishable tissue changes, from the standpoint of the pathologist working alone, producible by two different microorganisms (or by vitamin deficiencies, or by genetic mutations at two loci), they would be two different disease entities. So far as I am aware, this state of affairs is never strictly true. At least I have not come across any such in my reading of medicine, nor have my medical colleagues come up with any examples. The opposite case, of identical etiology (if etiology is identified with the specific etiological agents) but different pathology, is, of course, fairly common. Witness, for example, the numerous varieties of tuberculosis. The theoretical significance of a different bodily reaction to a particular invading organism is paralleled by great practical significance, since the physician does not treat tuberculous meningi-

tis, pulmonary tuberculosis, and tuberculous disease of the spine in precisely the same way. Of course when we expand the concept of etiology to mean both specific etiology and the predisposing, auxiliary, and precipitating causes (see, for example, Freud, 1895 as reprinted 1962), then the two different diagnoses can be separated (theoretically) into the same two taxa either on the basis of etiology or on the basis of pathology. Suppose that two patients' defensive reactions to invasion by an adequately infective number of microorganisms *Mycobacterium tuberculosis* do not succeed in preventing clinical involvement, but in one patient it takes the form of pulmonary tuberculosis and in the other patient the locus of tissue pathology is bone. In such an instance we must suppose that we have to deal with a *locus minoris resistentiae*, a disposition that must, strictly speaking, be counted as part of the "complete etiological equation." Hence a Utopian description of the etiological sequence as visualized by Omniscient Jones would distinguish the two cases just as clearly (specific etiology + dispositions of *locus minoris resistentiae*) as would the differential pathology (bone versus lung).

It is nevertheless a convenient simplification to distinguish pathology and etiology for many purposes, and I shall do so here. Figure 1 shows the situation in functional psychiatry by analogy to that in internal medicine. The diagram clarifies the core problem we face in setting up a reality-linked differential reinforcement schedule for the verbal behavior of participants in a clinical case conference. We do not *know* the pathology (character structure, psychodynamics, need/defense system, trait organization, basic temperamental parameters) of the patient; we only infer them, frequently with rather low degrees of probability and with marked disagreement among competent clinicians. But the situation is worse than it sounds. It is not merely a question (as it typically would be in internal medicine or neurology) about the *particulars* of the instant case, i.e., where this individual patient's pathology fits into the causal hyperspace of our received biochemistry, physiology, pathological anatomy, etc. In psychiatry there will be disagreements also about the nomothetic explanatory *system* that is admissible, to such an extent that at times there will be nearly zero overlap in the technical terminology between two clinicians. When we come to etiology, the situation is, if possible, worse still. One can find boarded psychologists and psychiatrists who believe that everyone is born with ab

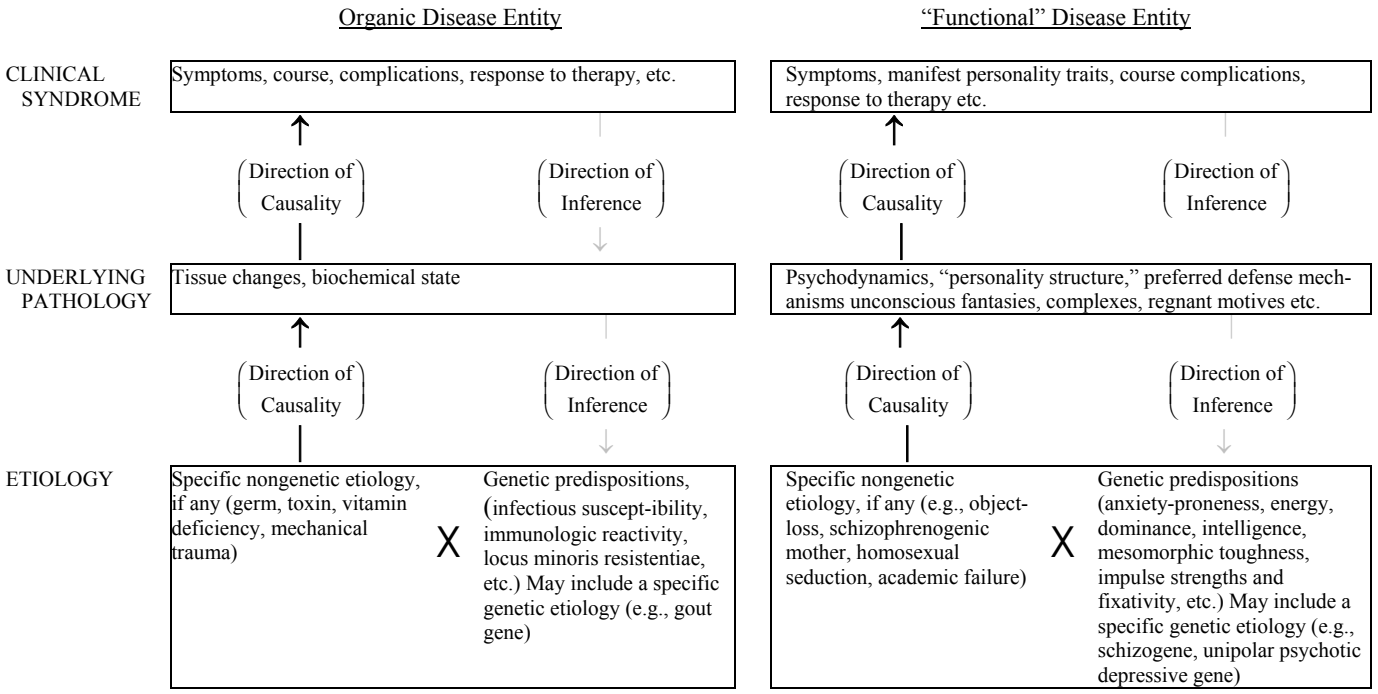

Figure 1. Clinical syndrome, underlying pathology, and etiology in organic disease entity and "functional" disease entity

solutely equal talent for developing schizophrenia (a position which I myself cannot see as now possible for a rational, informed mind); while others of equal professional qualifications (educational and experiential) may believe that both the occurrence and the form of schizophrenic disease lie wholly in genetic plus broadly constitutional factors, with psychological stresses and sociodynamics playing a negligible role. There is no Omniscient Jones psychopathologist whose biopsy report can stand as the umpire between such theoretical conflicts. Nobody can show slides demonstrating "superego lacunae." That fact, the absence of a definitive quasi-criterion, would seem to make insoluble the problem I am struggling with in this paper. Let us grant immediately that it is insoluble in the strict sense. But I want to argue that we can do considerably better than we have been, by adopting the unpopular medical model (with suitable adaptations to psychodiagnosis) and asking ourselves what would be the nearest equivalent to a pathologist's report.

The fundamental epistemological structure of the clinicopathological conference is easily characterized: It consists of withholding high-validity information, information that is quasi-definitive of the diagnosis, and requiring the participating clinicians to *infer* that high-validity, quasi-definitive information from other information which, at least in the average sense, possesses lower diagnostic validity. But this epistemic high validity is connected (as always) with the ontology, in this situation with the fact that the information in question is less remote in the causal chain than are the clinical symptoms, patient's complaints, response to treatment, hospital course, and so forth from the (definitive) pathological state cum etiological agent. That is, corresponding ontologically to causal closeness or intimacy (in some instances one could say "explicitly defined *meaning*") is something which, by virtue of its causal closeness, is also epistemologically "stronger" evidentially. In the limiting case, this epistemological strength is accepted as a criterion in the definitive sense mentioned above; although the extent to which this is true for the pathological examination of diseased tissue is easily exaggerated by psychologists. We withhold this high-validity information from the diagnosing clinician with the aim of sharpening his ability to make inferences from lower validity information, which he is often required to do in his clinical practice. Of course sometimes there is an artificiality about this in that we withhold information in the "gues-stimate" phase of the clinicopathological conference that the clinician in

his own practice might normally insist on having available before arriving at a decision. This element of artificiality is not considered too great a price to pay in order to attain the pedagogical aim of an objective criterion for differential reinforcement of inferential processes by a clinician diagnosing from presenting complaint, symptoms, signs, course, reaction to a therapeutic regime, and the like. I make this point because in searching for a realizable analogy to the epistemic circumstances of the clinicopathological conference when behavior disorder is the subject matter, one must be prepared for the objection that something artificial is being done. That is quite correct.

In addition to the epistemic factor of high validity deriving from the ontological factor of causal closeness, another influence tending to prevent case conferences in psychology or psychiatry from resembling a clinicopathological conference in medicine is the vagueness of the inferred statements, quite apart from the difficulty of ascertaining their truth. In a clinicopathological conference I might hazard the inference that the patient had an olfactory groove meningioma and I might be disappointed to learn that none such was found at autopsy, or that the patient was not operated on and the family refused permission for postmortem studies. The falsification of my inference, or the practical impossibility of checking it, does not arise from vagueness in the *meaning* of the expression "olfactory groove meningioma." *If* the tissue were available to the pathologist, whether or not the patient had an olfactory groove meningioma would be a question answerable with 99 percent certainty, whereas if I say that the patient has superego lacuna or (to use a once-favorite Rorschacher inference) has intrapsychic ataxia or—to gore Meehl's ox in a spirit of objectivity—that the patient is somewhat anhedonic, these expressions do not designate an even semi-precise state of affairs ontologically and therefore do not have a precise condition for their warranted assertability.

The rules of the game are so loose in psychiatry that it is interesting to speculate how far out, either in terms of conceptual vagueness or evidentiary weakness, one would have to go before his brethren called him on it. My teacher Dr. Starke Hathaway once mentioned that he was having so much trouble in one of his seminars in getting the psychiatry residents to adopt a critical posture, toward either the received doctrine or his own iconoclastic verbal productions, that he was about to go in and propound some absolute nonsense about the influence of sunspots

in schizophrenia, just to see whether they would rise to debate him or would dutifully note it down. There is in fact some quantitative evidence on this point (Goulett, 1965, pp. 8-9; cf. Goldsmith and Mandell, 1969), and all of us in the field of psychopathology at times permit ourselves utterances that, while perhaps not *utterly* devoid of empirical content, come about as close to it as one can find outside of Hegelian metaphysics.

We can formulate the psychiatric case conference problem thus, laying down conditions aimed at improvement but not unrealistically perfectionistic: We wish to divide the classes of information available at the time of the conference into two categories, the first category being available to the participant clinicians during their assessment process (including the conference discussion) and the second category being withheld from them until the end of the conference, presentation of this latter category of information being the differential reinforcement. In order for that division of information to serve its pedagogical function, we must meet three conditions:

1. The withheld information must be such that it will become reasonably clear ("objective") whether the statements inferred during the guesstimate phase of the conference are confirmed or refuted.

2. By and large, the statements belonging to the corpus of information withheld should themselves have an epistemologically privileged status in terms of the ontological structure, i.e., they should, in some sense that is defensible over the long run of patients, be closer to the underlying psychopathology/etiology than the evidentiary statements that are available in the guesstimate phase. While they cannot hope to have the status of the pathologist's report on a piece of biopsied or autopsied tissue, they should be analogous to it in the sense of being closer to that intrapsychic state of affairs that is nomologically definitive of the diagnostic entity or psychodynamic state.

3. This division of information on grounds of clearness and privileged evidentiary status must not do excessive violence to the ordinary clinical context. We are treating the participant clinicians as organisms whose behavior is being shaped up; we want to train them to do what they are going to do. Therefore while, in the interest of sharpening diagnostic skills by differential reinforcement, we may withhold some data that would normally be available at a comparable stage in the clinician's own

practice, the situation must not be so unlike that of ordinary clinical decision making as to be highly unrealistic.

Prima facie there are three sources of data theoretically capable of satisfying the first two of these conditions. First, we have the diagnostic, historical, and dynamic conclusions of the patient's therapist. In spite of a distressingly large element of subjectivity, these have merit in that they are based upon a larger sample of the patient's talk, gestures, fluctuations over time, response to probing, etc., than we have available when he is presented at the conference; and they are—unfortunately tied to the subjectivity—likely to be somewhat superior in quality to what we get in the case conference. But suppose it were seriously argued, as some hardnosed skeptics of my acquaintance would be willing to argue, that ten hours of psychotherapy is nearly worthless as a criterion of the truth about the patient's psychopathology. I cannot refute this skepticism. But then, by the same token, one would, in consistency, have to dismiss the conventional case conference enterprise in this field as fruitless. After all, if we think that *nothing* can be learned by observing and listening to a person talk and act in an interview, it is pointless to bring him into the conference to be interviewed by even the most able member of the professional staff. Second, we have the patient's behavior on the ward as observed by the ward personnel. It is a fairly objective fact that the patient refuses to take his medication, that he frequently approaches the nurses' cage with some sort of complaint, that he does not interact with other patients, that he sleeps soundly, and so forth. Third, whatever their intrinsic validity, the patient's psychometrics are a highly objective distillation of his responses to standard stimuli.

Each of these three information domains, which I shall label simply as "therapist ratings," "ward behavior," and "psychometrics," is likely to be qualitatively and quantitatively superior to what we can gather in the case conference. Further, it can be argued that to the extent that they cohere, they represent the closest we can come to a psychopathological equivalent of the pathologist's report in internal medicine. It is true that, with the exception of the psychometrics, the usual form in which these three data sources are available leaves much to be desired in the way of objectivity. But they do not have to be in the usual form, and part of my positive proposal is to modify that usual form in the direction of greater objectivity.

Consider first the psychotherapist's evaluation as a quasi-criterion. In order to reduce its vagueness, we require the psychotherapist to record his judgment in a standard format, such as the MHPA (Glueck, Meehl, Schofield, and Clyde, *Doctor's Sub-set, Forty Factors*, n.d.; Glueck, Meehl, Schofield, and Clyde, 1964; Glueck and Stroebel, 1969; Hedberg, Houck, and Glueck, 1971; Meehl, Lykken, Schofield, and Tellegen, 1971; Meehl, Schofield, Glueck, Studdiford, Hastings, Hathaway, and Clyde, 1962; Melrose, Stroebel, and Glueck, 1970; Mirabile, Houck, and Glueck, 1971). On the basis of such interviews, the psychotherapist has rated the patient on phenotypic variables (relatively close to behavior summaries), and the computer draws a factor profile. The same can be done for genotypic inferences by the therapist, although at present such profiles have not been developed for the MHPA genotypic pool. The obvious objection to taking this as a quasi-criterion is that although the therapist will have had a kind of clinical contact that is qualitatively superior to what we get at the case conference, and quantitatively he has had more hours than we have available, he may still be wrong, in the eyes of Omniscient Jones. There is no definitive answer to this objection, which is why I label it *quasi*-criterion. The best solution presently available, in my opinion, is to obtain independent ratings from a second skilled clinician who listens to tape recordings of the first clinician interviewing the patient. This suggestion may strike some readers as unrealistically burdensome for the staff, but it is not really so. There is evidence (Meehl, 1960—Chapter 6 above) to suggest that a psychotherapeutic interviewer's ratings converge rapidly toward the ratings he will be making after twenty-four hours of clinical contact, so that it would not usually be necessary for the second rating clinician to hear more than, say, two to four sessions of interviewing for present purposes. If this clinical job were performed solely for the purposes of the staff conference, it would be justifiable in the interest of better training; but of course carrying out such a rating task is itself a professional learning experience for the second judge and so can be defended on those grounds as well.

What do we do if the Q correlation between these two raters is low? The answer seems obvious to me: For most of our clinical case conferences, we would deliberately select those patients on whom there is satisfactory agreement between the two judges. Especially useful peda-

gogically would be patients on whom the two judges come to agree (convergence over time with more information) after poor agreement initially, presumably "hard" cases but not too hard to permit convergence given sufficient information.

From time to time, we would hold a staff conference on a patient where there was marked disagreement. In what follows I set aside that case and confine myself to the case in which there is a satisfactorily high Q correlation between the two independent raters. Some psychologists would argue that one should not infer validity from reliability, but this flat statement is misleading in some contexts. Reliability cannot prove validity, but it sometimes tends to support it. I urge that it does here—unless the enterprise is fruitless. (Nonpsychological example: If two surveying students independently come up with answers that a certain water tower is 847 feet from Stone *X* and is 200 feet high, this *does* tend to support the validity claim that these numbers are correct.) Ideally, if one were to set up a long-term program of improving the case conference along the lines suggested, it would be desirable to have a larger group of raters listening to the tapes or, better, a second interviewer making independent judgments on the basis of his own interview stimuli, together with a number of other raters listening to the tapes of this interviewer and the psychotherapist, Q correlating these ratings and arriving at an optimal statistical weight to be assigned. That is, we "calibrate" the (modal) therapist and (modal) tape listener on the basis of a larger number of expert clinicians, and use those statistical weights in future practice. The "best estimate" of the patient's characteristics is then a weighted sum of his therapist's judgments and the tape listener's judgments.

Predicting an MMPI profile seems like a rather silly thing to do, but it is really not. However, it is more realistic to predict not the profile itself but the modal Q-sort description of persons having the profile produced by the instant case. In order to bring this into coordination with the clinician's judgment, one must prepare an actuarial table such as the Marks-Seeman atlas (Marks and Seeman, 1963; Meehl, 1972b; Dahlstrom, Welsh, and Dahlstrom, 1972, pp. 307-339; Manning, 1971; Gilberstadt and Duker, 1965; Gilberstadt, 1970, 1972) and—again speaking ideally—one would want a large-scale investigation in which the MMPI-based description, the therapist's description, and the tape

listener's descriptions were thrown into one statistical hopper to yield weights for the best available construct-valid characterization of patients.

Finally, the traditional nurses' notes in the chart and the informal comments of nurses and psychiatric aides are poor substitutes for a checklist or rating scale as a summary of the patient's ward behavior (Glueck and Stroebel, 1969; Rosenberg, Glueck, and Stroebel, 1967).

Adopting the preceding suggestions would, one hopes, result in a set of high-construct-validity statements about the patient, with which statements made by clinicians participating in a staff conference could be compared. It is hardly feasible to require the conference participants to make a Q sort, but this does not present an insuperable difficulty for comparative purposes. What we do is to specify a set of domain rubrics for the characterization of patients, such as ego function, adequacy of control system, suitability for interpretative psychotherapy, acting-out tendencies, indications for this or that psychotropic medication, major mechanisms of defense, Murray needs, affective tone, insight, and nosological category. A mimeographed sheet could be passed out at the beginning of the conference to remind participants of these major sectors of patient description. For each descriptive area it will be possible to ascertain whether one has made inferences that correspond to those reached by an optimal statistical weighting of the nurses' observations of ward behavior, the psychotherapist's and the tape listener's ratings, and the personality description actuarially derived from the MMPI profile.

On the question of artificiality, there is admittedly something unrealistic about the proposed sequence of informational input. However, it is not as unrealistic as one might think at first—less so than the conventional case conference, in some ways. In clinical practice, for example, one does not normally have the psychometrics available to him at the time of his initial contact with the patient. He takes the history, does a Mental Status, and begins an inquiry into the patient's personality difficulties. Obviously he does not normally have any nurses' notes. If it is objected that we arrive at our integrated picture of the patient from all of the data, the answer is that we can do this after these quasi-criterial variables have been presented toward the end of the conference. Objections to blind diagnosis from personality tests seem to assume that one must choose whether to read the personality test in the light of other information or without it. The fact of the matter is, of course, that one

can read it both with and without the other evidence. (In my private office practice, and when I sit as a member of the State Hospital Review Board, I never look at MMPI profiles, if available, until *after* recording my interview impressions. The rationale is obvious: One set of "objective" numbers on a profile can infect my clinical impression, whereas infection going the other direction is impossible.) If it turns out that what the clinical staff concludes from the (selectively presented) life history plus the patient's behavior when interviewed in the conference is grossly out of line with the other data, this discrepancy should itself be a subject of discussion. The important point is that the inferences arrived at in the staff conference would include predictions about what the psychotherapists and the nurses and the MMPI said, and these agreements or discrepancies should constitute differential reinforcements for adequate versus inadequate clinical behavior by the participants. I may, of course, still think that I am right and that the MMPI is wrong; but it is a *fact* that I mispredicted the MMPI profile, or the personality profile based upon the MMPI.

While in this paper I am mainly concerned with the "intellectual" deficiencies that typically make clinical case conferences so boring, irksome, and educationally counterproductive, there are some practices of a procedural nature that help to make things dull, and need repair (whether or not my main suggestions for introducing a risky predictive element are acceptable to the reader). Since they are somewhat peripheral, I shall not develop argument at length but mainly list suggestions, with only the briefest summary of my reasons. Such a summary presentation will inevitably have a certain flavor of dogmatism about it.

When life-history material is presented (either initially, as in the currently accepted system, or later on, as in my suggested revision) it should be in documentary rather than oral form. It is preposterous that a roomful of highly paid faculty and busy psychiatric residents and graduate students in psychology should be forced to listen to somebody drone on about the fact that the patient's older brother died of appendicitis, that the patient had scarlet fever at age ten, that his uncle was a Swedenborgian, and the like. Even if the presentation of historical material were done more analytically and selectively than most residents or psychology trainees seem capable of doing, it would still be a terrible waste of time. There are certain kinds of basic "skeletal" data (geography, income, family's religion, organic illnesses, occupational history,

educational progress, and the like) which there is no justification for presenting orally to the group in the precious conference time period.

One of the vices of the present system is that so much time is thus spent (frequently because of inadequate preparation plus the inefficient oral presentation of history material) that in a conference scheduled to last an hour and a half, by the time we are ready to see the live patient, whoever is in charge of the conference and interviewing the patient is so uncomfortably conscious of how little time remains that the interview is almost pro forma. I have sat through conferences in which the first hour was spent in oral presentation of a melange of piddling and disconnected facts (including, say, that the patient had a great-uncle who died of cancer—the patient never having known his uncle); the patient then came in for ten minutes, leaving twenty minutes for a discussion of diagnosis, dynamics, and the treatment. This is simply absurd. For educational purposes (I am not here considering the sort of brief intake conference that many hospitals have on all new patients admitted since the previous morning) I think experience shows that no conference should be scheduled to run less than an hour and a half, and I myself would advocate two hours. Colleagues warn me that people's attention can't be held for two hours. I agree that you can't hold their attention with a bunch of poorly prepared third-raters doing a deadly presentation of meaningless material. But a variation in *who* is talking and *what* we are talking about, the difference between history and interview inputs, and especially the element of intellectual excitement generated by introducing elements of postdiction and prediction as I propose, should make it possible to hold people's interest for two hours. Most of us find we can run a two-hour seminar provided we run it right (that means, incidentally, not listening weekly to student literature reports!); and I therefore believe that two hours is feasible for good case conferences. Analogously to the seminar situation, anybody who has been around academia very long, and who remembers how he felt as a student, is aware that students do not much enjoy listening to each other. Admittedly student presentations serve an educational function for the presenter, although I see no reason for assuming that a first-year trainee, who has never attended any sort of conference before, is "ready" to begin his active learning process by presenting a case. In any event, we sacrifice a good deal of other students' valuable time when we force them to listen to an incompetent and boring presentation by somebody

who really isn't ready to do it. At the very least, I would suggest an alternation of major responsibility for presentation by advanced students, faculty, and near beginners. I recognize that this is predicated on the old-fashioned idea that full professors with twenty years of teaching, research, and clinical experience should, on the average, be capable of serving as educational models for fledgeling clinicians. Perhaps that is not true in psychiatry. If it is untrue, I think we ought to let the taxpayer in on the secret. If everybody is about equal in brains, skill, and knowledge, the taxpayer's elected representatives should be allowed to make up their minds whether they really want to pay Professor Fisbee $30,000 a year for functions like participating in a case conference, inasmuch as he doesn't know anything that a junior medical student or a first-year graduate in psychology doesn't know! (I have the impression that there is an economic question here, and conceivably an ethical one, but this is not the place to develop that line of thought.)

In either the present or the revised system, one must allow sufficient time so that discussion of the diagnosis, dynamics, and treatment can be carried on at a respectable level of intellectual depth. Questions like "Is there such an entity as schizophrenia?" or "Should the construct 'sociopathic personality' be defined mainly by psychological-trait criteria, or by life-history criteria (such as delinquency and underachivement)?" cannot be discussed meaningfully in five minutes. Many important questions which would presumably be part of the function of a clinical case conference are far better left undiscussed than discussed in a superficial, dilettante fashion. Nothing is more offensive to a first-class intellect than to have to listen to third-raters converse about an intrinsically fascinating and complicated topic. I have sat through case conferences in which nothing even moderately interesting took place until only ten minutes remained to discuss it. Of course this suggestion involves not merely suitable changes in the procedural aspects, and an enforcement of constraints on the consumption of time in certain ways by the participating personnel, but also more refractory problems, including the need for the power elite of a particular department to recognize that there *is* a scholarly and intellectually exciting way to discuss complicated subjects. Of course if someone does not have much of anything going on in his head, and suffers an impoverishment of mental furniture (common in the field we are discussing), he will not even understand why it is silly to discuss certain topics in ten minutes.

Part III: Concluding Remarks

This paper is a polemic. If some of my judgments seem harsh, I remind the reader that a psychiatric case conference involves the welfare of patients and their families, that we deal with the physical or psychological pain, the "success" or "failure," the incarceration or liberty, the economic dependency, and sometimes the life or death, of human beings for whom we have accepted some measure of responsibility. It will have been apparent that I am deeply offended by the intellectual mediocrity of what transpires in most case conferences; but this personal reaction is of only autobiographical interest. The ignorance, errors, scientific fallacies, clinical carelessness, and slovenly mental habits which I have discussed above are not merely offensive "academically." They have—sometimes dramatically—an adverse impact upon human lives. When a student therapist tells me that a patient he was treating went home on a weekend pass and blew his brains out, and I find out upon thorough exploration that this almost straight-A student (with high motivation and lofty ethical standards) *did not even know the patient's chart diagnosis*, I am not animated by sentiments of esteem or charity toward those responsible for this student's classroom instruction and clinical supervision. Furthermore, the taxpayer is shelling out some pretty fancy salaries for the professionals who conduct case conferences. One need not be a disciple of Ayn Rand to share her distaste for incompetence. I freely admit that a major component of my attack is a claim that the case conferences I have attended have been unrewarding to me largely because of the low level of competence—*both scientific and clinical*— of most participants.

But I hope to have said something more than this, something "constructive." I have tried to indicate that we face some special methodological difficulties in the psychiatric and psychological fields, difficulties so complicated and recalcitrant as to present major problems even for first-class scientists and practitioners. However, in order for those problems to be solved or ameliorated, it is first necessary to clean out the Augean stables—a thankless task, and one not calculated to win me any popularity contests. I have written bluntly and forcefully—no doubt some will think arrogantly—for which I herewith tender whatever apologies are due. I confess that I do not suffer fools gladly. But aside from the cathartic effect of writing this polemic, which expresses the

accumulated frustration and irritation of hundreds of hours of being subjected to this dismal business off and on for thirty years, before I quit entirely I cannot emphasize too strongly that part of the social and intellectual tradition of American psychiatry and clinical psychology tending to perpetuate the counterproductive mental habits described above is precisely this "buddy-buddy" syndrome which forbids anyone to call attention to instances of scientific or clinical incompetence, no matter how severe. So long as we operate on the principle that there are no standards of performance in this field, that everybody is equally bright, equally well read, equally skilled, equally logical, and equally experienced, Gresham's Law will, as usual, operate in the clinical case conference. There are too many psychoclinicians who implicitly equate the (valid) Popperian thesis that "Every informed, experienced, and intelligent professional is free to indulge his preferences among competing unrefuted conjectures" with the (preposterous) thesis that "Every professional or student is morally and intellectually entitled to persist in egregious mistakes, and it is wickedly authoritarian or snobbish to point them out." I take it that nobody who values the life of the intellect would subscribe to the latter thesis; and when it is applied in contexts involving psychological misery, physical health, economic dependency, crime, and sometimes death—as it is in the psychiatric case conference—such a maxim is not only foolish, it is downright immoral.

Finally, setting aside the unavoidable residuum of error inherent in the human condition, and the persistence of remediable errors among those professionals whose intellectual competence is simply not adequate to these difficult tasks, I have tried to offer at least the beginnings of a constructive plan for bringing the reinforcement schedule and cognitive feedback of the psychiatric case conference somewhat closer to those which prevail in the clinicopathological conference that has been so successful as a teaching device in the nonpsychiatric fields of medicine.

Addendum

As this volume was going to press, my psychiatric colleague Dr. Leonard L. Heston commented, on reading the manuscript that an alternative to the somewhat complicated construct-validity approach proposed herein as surrogate for clinicopathological conference criteria would be the

use of the follow-up. I am at a loss to understand my omitting this important alternative, except for the fact that my mental set was so strongly oriented toward solving the problem of *providing fairly quick diferential reinforcement*, of the kind that the internist receives at the end of each clinicopathological conference when the pathologist presents his quasi-criterial report on what the tissue showed. But, as Dr. Heston reminds me, we ought to be prepared to do some special things in psychiatry and clinical psychology, in trying to make up for the absence of the pathologist's report as a quasi-criterion of diagnosis. Dr. Heston points out that the clinician participating in a psychiatric case conference could be, so to speak, on record (we could even tape-record the conference—which might in itself tend to reduce some of the garbage generated!), and one's differential reinforcements would be forthcoming days, weeks, months (sometimes even years) later. Actually, there would be quite a few patients whose response to therapeutic intervention (e.g., phenothiazines in schizophrenia, electroplexy in psychotic depression, lithium carbonate in hypomania, valium in relatively uncomplicated anxiety states, RET in the "philosophical neurosis") would be ascertainable fairly soon after the case conference. Special provisions, including what might be a considerable financial outlay, would be necessary in order to achieve feedback on longer term forecasts. But I think that Dr. Heston's alternative suggestion is extremely important, and my discussion of the problem would be seriously defective without mention of it.

Of course, he and I agree that these are not really "competing alternatives," since both could be implemented, except insofar as we face the usual problem of opportunity costs. I have little doubt that the impact of some kinds of dramatic follow-up findings, their "convincing power," would be greater than the best souped-up, construct-valid, at-the-time quasi-criterion that could be devised with present methods. Two examples may be given.

Several years ago I had a two-hour diagnostic interview with a theology student from another city who presented with complaints of depression, anxiety, and "loss of interest," but who showed no clinical evidence of textbook schizophrenic thought disorder or markedly inappropriate affect. His flatness was no more severe on Mental Status appraisal than that which we find in many obsessional neurotics or other overintellectualizing, character-armored types. I daresay many of my

American colleagues, and the majority of European clinicians, would say that my interview-based diagnosis, "Schizotype, early stages of decompensation, marginal Hoch-Polatin syndrome," was an example of Meehl indulging his schizotypal hobby again. Nor would most such skeptics have been convinced—although they might have been somewhat influenced—by the (post-interview) scoring of the patient's MMPI profile, which yielded not merely the "gullwing curve" suggestive of pseudoneurotic schizophrenia but had a grossly psychotic (schizophrenic) configuration. As it happened, I subsequently found this patient to have shown up in a Canadian mental hospital with more obvious symptoms of schizophrenia; and then a year or so later, he again showed up (at the Minneapolis Veterans Administration Hospital) with symptoms of schizophrenia so unmistakable that even a very conservative diagnostician, such as Dr. Eliot Slater, would, I am sure, agree with the schizophrenic diagnosis there made.

A quicker but equally dramatic differential reinforcement for the diagnosticians I recall from my graduate school days, at a psychiatric grand rounds conducted by the late J. C. McKinley, M.D., co-author with Dr. Hathaway of the Minnesota Multiphasic Personality Inventory and then head of the Department of Neuropsychiatry. The patient seen in rounds that Saturday morning had presented with complaints of depression and anxiety, plus (as I recall it) rather vague nondelusional feelings that things seemed "not quite solid or real." He had a suspicious Rorschach with some rather bad 0^{-} responses but nothing so gross as confabulation or contamination, and with a marginal over-all form level; his MMPI was also borderline, although somewhat more in the psychotic than the neurotic direction by the then available "eyeballed" profile criteria. On interview a certain flatness, as in the preceding example, was clinically in evidence; but it was not gross and one could not really speak properly of Kraepelinian "inappropriate affect." After the interview was concluded and the patient had left the conference room a spirited debate took place among staff and students about whether the patient was an early schizophrenia or a neurotic with mixed anxiety, depression, and obsessional features. While we were still engaged in this debate (giving arguments pro and con from the history, the resident's Mental Status interview report, the interview that we had just observed, and the MMPI/Rorschach combination) the intern and charge nurse came back to inform us that the patient, after having left the conference

to be taken back to his room, had suddenly become mute and immobile, and was now standing in the corridor in a classical catatonic condition! This kind of quick and unmistakable feedback is of course unusual, but I don't think anybody who was present at that conference will ever forget the experience.

Allowing for the fact, as Jevons put it, that "Men mark where they hit and not when they miss," a series of such follow-up findings would either (a) show my colleagues that when I say somebody is a schizotype, I usually know what I am talking about or (b) convince me that I am erring in the direction of schizotypal overdiagnosis. On the other hand, I cannot close this necessarily brief discussion of Dr. Heston's proposed emphasis on follow-up as an alternative criterion without emphasizing that follow-up is unfortunately an asymmetrical affair, in the sense that certain *positive* subsequent developments are capable of strongly supporting some diagnoses as against others; but the theoretical and clinical positions with regard to "open-concept" entities like schizoidia, subclinical manic depression, and the like are such that the *failure* subsequently to develop unmistakable clinical phenomena pointing to diagnosis Dl and away from diagnosis D2 cannot, as is recognized by all sophisticated persons, be argued very strongly in the negative. (Cf. the diagnostic situation involving a patient at risk for Huntington's Disease, in a family strain with late onset, who shows irritability but no positive neurology at age 40, and dies of coronary disease two years later. Did he carry the Huntington gene? We will never know.) I regret that the limitations of space (in this already too long chapter) prevent my giving Dr. Heston's suggestion the full consideration that it merits.

I take this opportunity to add that since my scholarly psychiatric colleagues Drs. Leonard Heston and Neil Yorkston are now running a new weekly clinical case conference which is being inched up steadily to clinically and scientifically respectable standards, the title of this essay has become out-of-date for its author, since I am attending their conference regularly, with enjoyment and profit.

References

Alker, H. A., & H. G. Hermann. Are Bayesian decisions artificially intelligent? *Journal of Personality and Social Psychology*, 1971, 19, 31-41.

Beck, A. T. *Depression: Clinical, experimental and theoretical aspects*. New York: Harper and Row, 1967.

Beck, A. T., B. B. Sethi, & R. W. Tuthill. Childhood bereavement and adult depression. *Archives of General Psychiatry*, 1963, 9, 295-302.

Bergin, A., & H. Strupp. *Changing frontiers in the science of psychotherapy*. Chicago: Aldine, 1972.

Bleuler, E. *Dementia praecox; or, the group of schizophrenias* (1911), tr. Joseph Zinkin. New York: International Universities Press, 1950.

Burks, B. S., D. W. Jensen, & L. M. Terman. *Genetic studies of genius, 111: The promise of youth*. Stanford, Calif.: Stanford University Press, 1930.

Campbell, D. P. The vocational interests of American Psychological Association presidents. *American Psychologist*, 1965, 20, 636-644.

Campbell, D. P. *Handbook for the Strong Vocational Interest Blank*. Stanford Calif.: Stanford University Press, 1 7 .

Campbell, D. P., & D. W. Fiske. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 1959, 56, 81-105. Reprinted in E. Megargee, ed. *Research in clinical assessment*. New York: Harper and Row, 1966. Pp. 89-111. Also reprinted in W. Mehrens and R. L. Ebel, eds. *Principles of educational and psychological measurement*, Chicago: Rand McNally, 1967. Pp. 273-302. Also reprinted in D. N. Jackson and S. Messick, eds. *Problems in human assessment*. New York: McGraw-Hill, 1967. Pp. 124-132.

Castleman, B., & J. Burke. *Surgical clinicopathological conferences of the Massachusetts General Hospital*. Boston: Little, Brown, 1964.

Castleman, B., & H. R. Dudley, Jr. *Clinicopathological conferences of the Massachusetts General Hospital: Selected medical cases*. Boston: Little, Brown, 1960.

Castleman, B., & J. M. McNeill. *Bone and joint clinicopathological conferences of the Massachusetts General Hospital*. Boston: Little, Brown, 1967.

Castleman, B., & E. P. Richardson. *Neurologic clinicopathological conferences of the Massachusetts General Hospital*. Boston: Little, Brown, 1969.

Clark, K. E. *America's psychologists: A survey of a growing profession*, Washington, D.C.: American Psychological Association, 1957.

Cleckley, H. *The mask of sanity*. (4th ed.) St. Louis: C. V. Mosby, 1964.

Cronbach, L. J., & P. E. Meehl. Construct validity in psychological tests. *Psychological Bulletin*, 1955, 52, 281-302. Reprinted here as Chapter 1. Also available in the Bobbs-Merrill Reprint Series in the Social Sciences, no. P-82.

Dahlstrom, W. G., G. S. Welsh, & L. E. Dahlstrom. *An MMPI handbook, I: Clinical interpretation*. Minneapolis: University of Minnesota Press, 1972.

Dellas, M., & E. L. Gaier. Identification of creativity: The individual. *Psychological Bulletin*, 1970, 73, 55-73.

Dunnette, M. D. Use of the sugar pill by industrial psychologists. *American Psychologist*, 1957, 12, 223-225.

Einhorn, H. J. The use of nonlinear noncompensatory models in decision making. *Psychological Bulletin*, 1970, 73, 221-230.

Einhorn, H. J. Expert measurement and mechanical combination. *Organizational behavior and human performance*, 1972, 7, 86-106.

Erlenmeyer-Kimling, L., ed. Genetics and mental disorders. *International Journal of Mental Health*, 1972, 1, 1-230.

Fine, R. *The healing of the mind: The technique of psychoanalytic psychotherapy*. New York: McKay, 1971.

Forer, B. R. The fallacy of personal validation: A classroom demonstration of gullibility. *Journal of Abnormal and Social Psychology*, 1949, 44, 118-123.

Frank, G. H. The role of the family in the development of psychopathology. *Psychological Bulletin*, 1965, 64, 191-205.

Freud, S. A reply to criticisms on the anxiety-neurosis (1895). In *Collected papers*, I. London: Hogarth Press, 1950. Pp. 107-127. Also in J. Strachey, ed. *Standard edition of the complete psychological works of Sigmund Freud*, III. London: Hogarth Press, 1962. Pp. 119-139.

Gilberstadt, H. Comprehensive MMPI codebook for males. IB 11-5. Washington, D.C.: Veterans Administration, 1970.

Gilberstadt, H. *Supplementary MMPI codebook for VA male medical consultation*. IB 11-5, Supplement 1. Washington, D.C.: Veterans Administration. 1972

Gilberstadt, H., & J. Duker. *A handbook for clinical and actuarial MMPI interpretation*. Philadelphia: W. B. Saunders, 1965.

Glueck, B. C., Jr., P. E. Meehl, W. Schofield, & D. J. Clyde. *Minnesota-Hartford Personality Assay: Doctor's sub-set*. Hartford, Conn.: Institute of Living, n.d.

Glueck, B. C., Jr., P. E. Meehl, W. Schofield, & D. J. Clyde. *Minnesota- Hartford Personality Assay: Forty factors*. Hartford, Conn.: Institute of Living, n.d.

Glueck, B. C., Jr., P. E. Meehl, W. Schofield, & D. J. Clyde. The quantitative assessment of personality. *Comprehensive Psychiatry*, 1964, 5, 15-23.

Glueck, B. C., Jr., & C. F. Stroebel. The computer and the clinical decision process, 11. *American Journal of Psychiatry*, Supplement, 1969, 125, 2-7.

Goldberg, L. R. Simple models or simple processes? Some research on clinical judgments. *American Psychologist*, 1968, 23, 483-496.

Goldberg, L. R. Man versus model of man: A rationale, plus some evidence, for a method on improving on clinical inferences. *Psychological Bulletin*, 1970, 73, 422-432.

Goldsmith, S. R. and A. 1. Mandell. The psychodynamic formulation: A critique of a psychiatric ritual. *American Journal of Psychiatry*, 1969, 125, 1738-1743.

Gottesman, I. 1., & J. Shields. *Schizophrenia and genetics: A twin study vantage point*. New York: Academic Press, 1972.

Goulett, H. M. *The insanity defense in criminal trials*. St. Paul, Minn.: West Publishing Company, 1965.

Hacking, I. *Logic of statistical inference*. Cambridge: Cambridge University Press, 1965.

Hedberg, D. L., J. H. Houck, & B. C. Glueck, Jr. Tranylcypromine-trifluoperazine combination in the treatment of schizophrenia. *American Journal of Psychiatry*, 1971, 127, 1141-1146.

Heston, L. Genes and schizophrenia. In J. Mendels, ed. *Textbook of biological psychiatry*. New York: Wiley-Interscience, 1973 (in press)

Hoch, P., & P. Polatin. Pseudoneurotic forms of schizophrenia. *Psychiatric Quarterly*, 1949, 3, 248-276.

Hollingshead, A. B., & F. C. Redlich. *Social class and mental illness: A community study*. New York: Wiley, 1958.

Hollingworth, L. S. *Gifted children*. New York: Macmillan, 1926.

Kendall, M. G. On the reconciliation of theories of probability. *Biometrika*, 1949, 36,101-116.

Klebanoff, L. B. Parental attitudes of mothers of schizophrenic, brain-injured and retarded, and normal children. *American Journal of Orthopsychiatry*, 1959, 29, 445-454.

Kleinmuntz, B. *Clinical information processing by computer*. New York: Holt Rinehart, and Winston, 1969.

Kleinmuntz. B., ed. *Formal representation of human judgment*. New York: Wiley, 1968.

Levi, I. *Gambling with truth: An essay on induction and the aims of science*. New York: Knopf, 1967.

Livermore, J. M., C. P. Malmquist, and P. E. Meehl. On the justifications for civil commitment. *University of Pennsylvania Law Review*, 1968, 117, 75-96.

Loevinger, J. Objective tests as instruments of psychological theory. *Psychological Reports*, Monograph Supplement 9, 1957, 3, 635-694. Reprinted in D. N. Jackson and S. Messick, eds. *Problems in human assessment*. New York: McGraw-Hill, 1967.

Lundberg, G. A. Case-studies vs. statistical method—an issue based on misunderstanding. *Sociometry*, 1941, 4, 379-383.

McNemar, Q. Lost: Our intelligence? why? *American Psychologist*, 1964, 19, 87 1-882.

Malmquist, C. P. Depression and object loss in acute psychiatric admissions. *American Journal of Psychiatry*, 1970, 126, 1782-1787.

Manning, H. M. Programmed interpretation of the MMPI. *Journal of Personality Assessment*, 1971, 35, 162-176.

Manosevitz, M., G. Lindzey, and D. D. Thiessen. *Behavioral genetics: Method and research*. New York: Appleton-Century-Crofts, 1969.

Marks, P. A., and W. Seeman. *The actuarial descriplion of abnormal personality*. Baltimore: Williams and Wilkins, 1963.

Marks, P. A., and J. O. Sines. Methodological problems of cookbook construction. In J. Butcher, ed. *MMPI: Research developments and clinical applications*. New York: McGraw-Hill, 1969. Pp. 71-96.

Meehl, P. E. An examination of the treatment of stimulus patterning in Professor Hull's Principles of behavior. *Psychological Review*, 1945, 52, 324-332. (a)

Meehl, P. E. *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence*. Minneapolis: University of Minnesota Press, 1954. (a).

Meehl, P. E. Clinical versus actuarial prediction. *Proceedings of the 1955 Invitational Conference on Testing Problems*- Princeton, N.J.: Educational Testing Service, 1956. Pp. 136-141. (a)

Meehl, P. E. Symposium on clinical and statistical prediction. *Journal of Counseling Psychology*, 1956, 3, 163-173. (b)

Meehl, P. E. Wanted—a good cookbook. *American Psychologist*, 1956, 11, 263-272. Reprinted here as Chapter 3. (c)

Meehl, P. E. When shall we use our heads instead of the formula? *Journal of Counseling Psychology*, 1957 4, 268-273. Reprinted here as chapter 4 Available also in the Bobbs-Merrill Reprint Series in the Social Sciences, no P-SI9.

Meehl, P. E. A comparison of clinicians with five statistical methods of identifying psychotic MMPI profiles. *Journal of Counseling Psychology*, 1959, 6, 102-109. (a)

Meehl, P. E. Some ruminations on the validation of clinical procedures. *Canadian Journal of Psychology*, 1959, 13, 102-128. Reprinted here as Chapter 5. Also available in the Bobbs-Merrill Reprint Series in the Social Sciences, no. 517. (b)

Meehl, P. E. The cognitive activity of the clinician. *American Psychologist*, 1960, 15, 19-27. Reprinted here as Chapter 6. Also available in the Bobbs-Merrill Reprint Series in the Social Sciences, no. P-518.

Meehl, P. E. Schizotaxia, schizotypy, schizophrenia, *American Psychologist*, 1962, 17, 827-838. Reprinted here as Chapter 7. Also available in the Bobbs-Merrill Reprint Series in the Social Sciences, no. P-516.

Meehl, P. E. *Manual for use with checklist of schizotypic signs*. Minneapolis: Psychiatric Research Unit, University of Minnesota Medical School, 1964.

Meehl, P. E. Seer over sign: The first good example. *Journal of Experimental Research in Personality*, 1965, 1, 27-32. (c)

Meehl, P. E. What can the clinician do well? In D. N. Jackson and S. Messick, eds. Problems in human assessmenr. New York: McGraw-Hill, 1967. Pp. 594-599. Reprinted here as Chapter 9. (b)

Meehl, P. E. Nuisance variables and the ex post facto design. *Reports from the Research Laboratories of the Department of Psychiatry, University of Minnesota*. Report no. PR-69-4. Minneapolis: University of Minnesota, April 15, 1969. (b)

Meehl, P. E. Nuisance variables and the ex post facto design. In M. Radner and S. Winokur, eds. *Minnesota studies in the philosophy of science*, IV. Minneapolis: University of Minnesota Press, 1970. (Expanded version of Meehl, 1969b.) Pp. 373-402. (a)

Meehl, P. E. Psychology and the criminal law. *University of Richmond Law Review*, 1970, 5, 1-30. (b)

Meehl, P. E. Some methodological reflections on the difficulties of psychoanalytic research. In M. Radner and S. Winokur, eds. *Minnesota studies in the philosophy of science*, IV. Minneapolis: University of Minnesota Press, 1970. Pp. 403-416. (c)

Meehl, P. E. Reactions, reflections, projections. In J. Butcher, ed. *Objective personality assessment: Changing perspectives*. New York: Academic Press, 1972. Pp. 131-189. (b)

Meehl, P. E. Specific genetic etiology, psychodynamics and therapeutic nihilism. *International Journal of Mental Health*, 1972, 1, 10-27. Reprinted here as Chapter 11. (c)

Meehl, P. E. MAXCOV-HITMAX: A taxonomic search method for loose genetic syndromes. 1973. Printed here as Chapter 12.

Meehl, P. E. The concept 'specific etiology': Some quantitative meanings. To appear.

Meehl, P. E., and W. G. Dahlstrom. Objective configural rules for discriminating psychotic from neurotic MMPI profiles. *Journal of Consulting Psychology*, 1960, 24, 375-387.

Meehl, P. E., D. T. Lykken, W. Schofield, and A. Tellegen. Recaptured-item technique (RIT): A method for reducing somewhat the subjective element in factor-naming. *Journal of Experimental Research in Personality*, 1971, 5, 171-190.

Meehl, P. E., and A. Rosen. Antecedent probability and the efficiency of psychometric signs, patterns, or cutting scores. *Psychological Bulletin*, 1955, 52, 194-216. Reprinted here as Chapter 2. Also available in the Bobbs-Merrill Reprint Series in the Social Sciences, no. P-514.

Meehl, P. E., W. Schofield, B. C. Glueck, Jr., W. B. Studdiford, D. W. Hastings, S. R. Hathaway, and D. J. Clyde. *Minnesota-Ford pool of phenotypic personality items*. (August 1962 ed.) Minneapolis: University of Minnesota, 1962.

Melrose, J. P., C. F. Stroebel, and B. C. Glueck, Jr. Diagnosis of psychopathology using stepwise multiple discriminant analysis, I. *Comprehensive Psychiatry*, 1970, 11, 43-50.

Meltzoff, J., and M. Kornreich. *Research in psychotherapy*. New York: Atherton Press, 1970.

Mirabile, C. S., J. H. Houck, and B- C- Glueck, Jr. Computer prediction of treatment success. *Comprehensive Psychiatry*, 1971, 12, 48-53.

Morrison, D. E., and R. E. Henkel, eds- *The significance test controversy*. Chicago: Aldine, 1970.

Myers, J. K., and L. Schaffer. Social stratification and psychiatric practice. *American Sociological Review*, 1954, 19, 307-310.

Nash, L. K. *The atomic-molecular theory*. Cambridge, Mass.: Harvard University Press, 1950.

Pankoff, L. D., and H. B. Roberts. Bayesian synthesis of clinical and statistical prediction. *Psychological Bulletin*, 1968, 70, 762-773.

Paterson, D. G. Character reading at sight of Mr. X according to the system of Mr. P. T. Barnum. Unpublished, mimeographed. First printed in M. L. Blum and B. Balinsky. *Counseling and psychology*. New York: Prentice-Hall, 1951. P. 47. Reprinted in M. D. Dunnette. Use of the sugar pill by industrial psychologists. *American Psychologist*, 1957, 12, 223.

Raiffa, H. *Decision analysis: Introductory lectures on choices under uncertainty*. Reading, Mass.: Addison-Wesley, 1968.

Rosen, A. Detection of suicidal patients: An example of some limitations in the prediction of infrequent events. *Journal of Consulting Psychology*, 1954, 18, 397-403.

Rosen, E., R. E. Fox, and I. Gregory. *Abnormal psychology*. (2nd ed.) Philadelphia: W. B. Saunders, 1972.

Rosenberg, M., B. C. Glueck, Jr., and C. F. Stroebel. The computer and the clinical decision process. *American Journal of Psychiatry*, 1967, 124, 595-599.

Rosenthal, D. *Genetic theory and abnormal behavior*. New York: McGraw-Hill, 1970.

Sarbin, T. R. A contribution to the study of actuarial and individual methods of prediction. *American Journal of Sociology*, 1942, 48, 593-602.

Savage, L. J. *The foundations of statistics*. New York: Wiley, 1954.

Sawyer, J. Measurement and prediction, clinical and statistical. *Psychological Bulletin*, 1966, 66, 178-200.

Schmidt, H. O. and C. P. Fonda. The reliability of psychiatric diagnosis: A new look. *Journal of Abnormal and Social Psychology*, 1956, 52 262-267.

Schofield, W. *Psychotherapy: The purchase of friendship*. Englewood Cliffs, N.J.: Prentice-Hall, 1964.

Schofield, W., and L. Balian. A comparative study of the personal histories of schizophrenic and nonpsychiatric patients. *Journal of Abnormal and Social Psychology*, 1959, 59, 216-225.

Shaffer, L. F. Of whose reality I cannot doubt. *American Psychologist*, 1953, 8, 608-623.

Strong, E. K., Jr. *Vocational interests of men and women*. Stanford, Calif.: Stanford University Press, 1943.

Sundberg, N. D. The acceptability of "fake" Versus "bona fide" personality test interpretations. *Journal of Abnormal and Social Psychology*, 1955, 50, 145-1 17

Tallent, N. On individualizing the psychologist's clinical evaluation. *Journal of Clinical Psychology*, 1958, 14, 243-244.

Terman, L. M. *Genetic studies of genius, 1: Mental and physical traits of a thousand gifted children*. Stanford, Calif., Stanford University Press, 1925.

Terman, L. M., and M. H. Oden. *Genetic studies of genius, IV: The gifted child grows up*. Stanford, Calif.: Stanford University Press, 1947.

Terman, L. M., and M. H. Oden. *Genetic studies of genius, V: The gifted group at mid-life*. Stanford, Calif.: Stanford University Press, 1959.

Thorndike, R. L. The psychological value systems of psychologists. *American Psychologist*, 1954, 9, 787-789.

Thorndike, R. L The structure of preferences for psychological activities among psychologists. *American Psychologist*, 1955, 10, 205-207.

Ulrich, R. E., T. J. Stachnik, and N. R. Stainton. Student acceptance of generalized personality interpretations. *Psychological Reports*, 1963, 13, 831-834.

pdf by ljy August 2002