JOURNAL OF PERSONALITY ASSESSMENT, 1986, 50(3), 370-375.

Causes and Effects of My Disturbing Little Book

Paul E. Meehl

University of Minnesota

Review and reflection indicate that no more than 5% of what was written in the 1954 book entitled, *Clinical Versus Statistical Prediction* (Meehl, 1954), needs to be retracted 30 years later. If anything, these retractions would result in the book's being more actuarial than it was. Seven factors appear to account for the failure of mental health professionals to apply in practice the strong and clearly supported empirical generalizations demonstrating the superiority of actuarial over clinical prediction.

In my early teens, I had intended to be a lawyer, but a copy of Karl Menninger's The *Human Mind* (1930) fell into my hands. Because it was effective dramatic bibliotherapy for me, I decided overnight to change my vocational goal to psychotherapist. Having reread Menninger's book recently, I still think it is one of the best expositions of basic psychodynamics. Minnesota's Psychology and Psychiatry Departments were strongly anti-Freudian, the psychologists being behavioristic in theoretical orientation and (except for Fred Skinner) highly statistical in method. The Neuropsychiatry Department stressed neurology, and although excellent in descriptive psychiatry, it was weak in psychodynamics and psychotherapy. My undergraduate minor was biometry (as was true of most Minnesota psychology majors in those days). Although my doctoral advisor, Starke R. Hathaway, inventor of the Minnesota Multiphasic Personality Inventory (MMPI), put much greater emphasis on clinical know-how than the more "academic" professors on the nonmedical side of the campus, he stressed clinical observation of the descriptive nosology kind, minimally inferential. The only "Freudian" component in his thinking was a strong emphasis on spotting the patient's preferred mechanisms of defense.

You can easily see how a bright, thoughtful student, drawn to psychology via Menninger's book (and I had read a great deal of Freud, Jones, and Ferenczi in high school), confronted with this psychometric, behavioristic, anti-Freudian intellectual environment, would be forced to reflect on the relationship between these two very different approaches. There was the further contribution from philosophy of science because one of the great Vienna positivists, the author of the first article in English on that position and the inventor of the phrase *logical positivism,* Herbert Feigl, came to the Minnesota Philosophy Department in my senior year. The Department of Sociology was chaired by Stuart Chapin, a leading expounder of quantitative sociology. George Lundberg had been at Minnesota, and a criminologist named Vold, who worked on actuarial prediction of parole violation, was still on the faculty when I was a student. Finally, at the time I was taking Donald G. Paterson's famous course in differential psychology, Paterson was engaged in an extended controversy, partly in the literature but mostly by correspondence, with Gordon Allport on the idiographic–nomothetic business. Arthur H. Brayfield, who later served our association as executive officer, was completing his doctoral work under Paterson in the mid-1940s, and it was Brayfield who first called my attention to Dr. Sarbin's (1943) classic article in the *American Journal of Sociology.*

As a result of reading Sarbin, I was determined to write a paper on the clinical–actuarial question, but I found that I was uncertain about exactly where I stood. The more I thought about it and the more talked to scholars about it, the more difficult and complicated it seemed and the longer the manuscript became. Some time around 1950, I realized I had a small book, which I was unable to place with any publisher (including two publishers with whom I had an intimate connection) because they said it would not sell! It has now gone through seven printings. From the time I read Sarbin to the appearance of my book in 1954, there was about a 10-year period in which a sizable portion of my intellectual effort was devoted to reading, reflection, and discussion of this question. False modesty aside, when a fairly bright fellow like me, with training in philosophy of science, mathematical statistics, and clinical psychology, and with intense motivations (originating as I have described), sympathetic to both the clinical and actuarial approaches, thinks about a question for a decade, what he comes out with is likely to be pretty good. I find, as a result, that I need to retract at most 5% of what I wrote in the first edition. Further, to the extent that I have to retract, I am afraid that I have to be more actuarial than the book was because my main proclinical considerations partly relied on mathematical patterning effects (which turn out to be rare, unstable, hard to detect, and of low incremental validity when present) and—what I emphasized most—the kind of complicated psychodynamic inference, with heavy idiographic component, that occurs in psychoanalysis, which seems less important today.

The reviews were uniformly favorable, even from clinicians who were upset, as many clinicians were. I am told that half the clinical faculty at one large Freudian oriented midwest university were plunged into a 6-month reactive depression as a result of my little book. Despite some allegations to the contrary, I contend that I was scrupulously fair to the clinician, as is understandable given the motivations that impelled me to write it in the first place. I have not done a formal content analysis, but it is clear that by far the larger part of the discussion is devoted to explaining what I thought was mistaken about the Sarbin–Lundberg position rather than what was valid in it. The empirical chapter, summarizing the available 22 studies, I viewed as the least important part of the book. I was distressed to see that some clinicians reacted as if that was the only chapter I had written.

Some critics asked a question that Dr. Holt still asks, and which I confess I am totally unable to understand: Why should Sarbin and Meehl be fomenting this needless controversy? Let me state as loudly and as clearly as I can manage, even if it distresses people who fear disagreement, that Sarbin and I did not artificially concoct a controversy or foment a needless fracas between two methods that complement each other and work together harmoniously. I think this is a ridiculous position when the context is the pragmatic context of decision making. You have two quite different procedures for combining a finite set of information to arrive at a predictive decision. It is obvious from the armchair, even if the data did not show it overwhelmingly, that the results of applying these two different techniques to the same data set do not always agree. On the contrary, they disagree a sizable fraction of the time. Now if a four-variable regression equation or a Glueck actuarial table tells the criminal court judge that this particular delinquent will probably commit another felony in the next .3 years and if a case conference or a social worker says that he will probably not, it is absurd to say that Sarbin and I have "fomented a controversy" about how the judge should proceed. The plain fact is that he cannot act in accordance with both of these incompatible predictions.

Nobody disputes that it is possible to improve clinicians' practices by informing them of their track records actuarially. Nobody has ever disputed that the actuary would be well advised to listen to clinicians in setting up the set of variables.

Why should people have been so surprised by the empirical results in my summary chapter? Surely we all know that the human brain is poor at weighting and computing. When you check out at a supermarket, you don't eyeball the heap of purchases and say to the clerk, "Well it looks to me as if it's about $17.00 worth; what do you think?" The clerk adds it up. There are no strong arguments, from the armchair or from empirical studies of cognitive psychology, for believing that human beings can assign optimal weights in equations subjectively or that they apply their own weights consistently, the query from which Lew Goldberg derived such fascinating and fundamental results.

It is true that there are three things that even the modern super computer cannot as yet do very well, and I have no theoretical stake whatsoever in disputing that. The first is pattern recognition, where the computer still cannot recognize a slightly battered typeface as well as the human eye and brain; second is the translation of languages; and third is the construction of theories. That these three tasks are still beyond the computer's talents gives the tip-off as to why my favorite nonactuarial examples in the book, and in subsequent publications, are all taken from psychoanalytic therapy. However, although still qualitatively correct, this psychoanalytic therapy point is far less important quantitatively than it was 30 years ago, for three present reasons and one likely future development. First, contemporary psychoanalytic therapy is rarely classical and lays far less emphasis on interpretation of unconscious thematic content and historical constructions than in Freud's day or even when I was an analysand in the 1950s. There is, for instance, less dream interpretation or reliance on free association as a technical device. Many analysts no longer even impose the fundamental rule. Second, there is more evidence of relatively poor interanalyst reliability of inferences from the same dream, the same associations, or the same projective productions. Third, the efficacy of psychoanalysis as a treatment mode is now known to be relatively poor compared to behavior modification and rational emotive therapy. It is possible for a sane, rational, informed person to doubt that psychoanalysis, in anything like the form that emphasizes complex inferences to unconscious processes, is an effective procedure at all, although it is certainly a fascinating and educational growth experience as I continue to believe was true in my case.

For the future, because computers today can play master chess, have sometimes defeated master players, and have even begun to do primitive low-order theory construction when presented with a data set, it is rash indeed to say that we will never be able to write software for interpreting dreams, with a reliability as good as that of the human dream analyst.

I repeat that I am not retracting the qualitative points I made about the unique functions of the clinical brain in my book; I am simply saying that their quantitative importance in clinical practice has markedly declined in the ensuing 30 years.

However, the biggest factor on which Dr. Holt has repeatedly agreed with me in his writings and in correspondence, and which no informed person who knows about firing-line clinical predictions could dispute, is that 95% of the ordinary decisions made by working practitioners, whether psychiatrist, psychologist, or social worker, are not comparable in richness and subtlety to that of a good psychoanalytic hour. The special function of the skilled clinical brain that I was at such pains to emphasize against Sarbin and Lundberg rarely operates in the ordinary workaday predictions of a parole board or in forecasting whether somebody will do well in law school, or respond to Elavil, or continue in group therapy.

In order to use theoretical concepts fruitfully in making predictions for concrete cases, one requires a well-corroborated theory, which has high verisimilitude and includes almost all of the relevant variables, and an accurate technology of measurement, including access to the initial and boundary conditions of the system to be predicted and negligible influence of what Paul Horst called "contingency factors." None of these conditions is met in our routine clinical forecasting situation. As regards contingency factors, the interesting thing about them is that they are in a funny sense included in the actuarial prediction procedure in that the weights and residuals take them collectively into account, the known and unknown alike. The theory-based prediction does not do this.

I would be interested in knowing whether any readers want to controvert the following claim: There is no controversy in social science that shows such a large body of qualitatively diverse studies coming out so uniformly in the same direction as this one. When you are pushing 90 investigations, predicting everything from the outcome of football games to the diagnosis of liver disease and when you can hardly come up with a half dozen studies showing even a weak tendency in favor of the clinician, it is time to draw a practical conclusion, whatever theoretical differences may still be disputed. Why, then, is such a strongly and clearly supported empirical generalization not applied in practice, particularly because there are no plausible theoretical reasons to have expected otherwise in the first place? Not to argue ad hominem but to explain after the fact, I think this is just one more of the numerous examples of the ubiquity and recalcitrance of irrationality in the conduct of human affairs. If I had to spell out the psychosocial contributors to this irrationality in the present situation, I would list seven factors as follows:

1. Sheer ignorance: It amazes me how many psychologists, sociologists, and social workers do not know the data, do not know the mathematics and statistics that are relevant, do not know the philosophy of science, and are not even aware that a controversy exists in the scholarly literature. But what can you expect, when I find that the majority of clinical psychology trainees getting a PhD at the University of Minnesota do not know what Bayes' Theorem is, or why it bears upon clinical decision making, and never heard of the Spearman–Brown Prophecy Formula!

2. The threat of technological unemployment: If PhD psychologists spend half their time giving Rorschachs and talking about them in team meetings, they do not like to think that a person with an MA in biometry could do a better job at many of the predictive tasks.

3. Self-concept: "This is what I do; this is the kind of professional I am." Denting this self-image is something that would trouble any of us, quite apart from the pocketbook nerve.

4. Theoretical identifications: "I'm a Freudian, although I have to admit Freudian theory doesn't enable me to predict anything of practical importance about the patients." Although not self-contradictory, such a cognitive position would make most of us uncomfortable.

5. Dehumanizing flavor: Somehow, using an equation to forecast a person's actions is treating the individual like a white rat or an inanimate object, as an *it* rather than as a *thou;* hence, it is spiritually disreputable.

6. Mistaken conceptions of ethics: I agree with Aquinas that *caritas* is not an affair of the feelings but a matter of the rationally informed will. If I try to forecast something important about a college student, or a criminal, or a depressed patient by inefficient rather than efficient means, meanwhile charging this person or the taxpayer 10 times as much money as

I would need to achieve greater predictive accuracy, that is not a sound ethical practice. That it feels better, warmer, and cuddlier to me as predictor is a shabby excuse indeed.

7. Computer phobia: There is a kind of general resentment, found in some social scientists but especially people in the humanities, about the very idea that a computer can do things better than the human mind. I can detect this in myself as regards psychoanalytic inference and theory construction, but I view it as an irrational thought, which I should attempt to conquer.

People sometimes ask me whether I am disappointed by the relatively feeble impact of that book and of the many studies that were stimulated by it. The only sizable influence has been the proliferation of MMPI cookbook interpretations, in excess of their demonstrated validity. I suppose I am mildly disappointed, but you must remember that by 1954, after 10 years of corresponding, clinical conversation, and classroom teaching, I had quite a good inductive base to conclude how clinical psychologists react to this matter. Hence, I really didn't expect people to be able to think rationally about it, despite my heroic struggles to do so. As a 60% disciple of Albert Ellis, I do not view human irrationality as confined to mentally ill patients or even to the milder maladjustments we see in outpatient psychotherapy, but rather as par for the course, as fairly standard for the human condition. I accept the scholar's role of "clerk" in Julian Benda's sense, and I hold that the clerk's social function is to think as clearly and objectively as human frailty permits and to publish the deliverances of his or her cerebration in the most effective way possible. But the scholar, the professional intellectual—unlike the preacher, politician, or advertising agent—is not mainly in the public relations business. When a scholar has done a responsible job of thinking and communicating his or her thoughts, the task is at an end. Thus, I have learned to develop a certain Buddhistic detachment about the matter. Suppose a social worker confidently tells me that of course we can predict how this delinquent will do on probation by reflecting on psychodynamic inferences and subjective impressions, recorded in a 10-page presentence investigation, despite the malignant rap sheet record and acting-out psychometrics, and the officer's comment that "he's a real mean, tough street kid." Well, I remind myself that Omniscient Jones has not put me in charge of reforming the world.

REFERENCES

Meehl, P. E. (1954). *Clinical versus statistical prediction.* Minneapolis: University of Minnesota Press.

Menninger, K. A. (1930). *The human mind.* New York: Literary Guild of America.

Sarbin, T. R. (1943). A contribution to the study of actuarial and individual methods of prediction. *American Journal of Sociology,* 48, 593-602.

pdf by LJY
January 2015