

The Role of Noncognitive Measures in Medical School Admissions

Lawrence J. Stricker
Educational Testing Service

The value of noncognitive measures in medical school admissions was assessed in light of the existing literature. These measures appear to have limited usefulness in predicting success in academic work but may be valuable in forecasting both performance in clinical training and performance as a physician, as well as forecasting choice of the type of practice and its location. Noncognitive measures are useful as predictors of such criteria and may be valuable in forecasting the decisions of admissions committees; their use as moderator variables, however, is problematic. Newer personality and interest inventories, along with biographical questionnaires, are the most promising measures. Older interest inventories may have some value; but traditional personality inventories and projective techniques, as well as interviews, seem to have limited usefulness. The merit of the other measures is uncertain: Letters of recommendation are probably of little use; but cognitive style tests, objective performance devices, and special adaptations of projective techniques deserve more attention. The evaluation of noncognitive measures is hampered by inadequate criteria. Distortion by examinees threatens all self-report measures but can be controlled.

The admissions process in medical schools, as in most academic institutions, has traditionally relied heavily on cognitive variables, such as ability tests and course grades (Cuca, Sakakenny, & Johnson, 1976; Hutchins, 1977). This

emphasis is understandable, for it is obviously critical that students have the intellectual capacity to complete their studies and to function as physicians. Despite this cognitive focus, it has also been recognized that noncognitive variables, which predominantly reflect personality aspects of individual differences (Messick, 1979), are relevant, as attested to by the routine use of letters of recommendations and interviews (Cuca et al., 1976). Interest in the use of measures of noncognitive characteristics in admissions has heightened in recent years as a result of (1) dramatic changes in the nature of the population of medical school applicants (the size and quality of this group at present is such that students with the ability to graduate can be selected with little difficulty) and (2) increased concern about the careers of those who are being graduated (the quality of care that these graduates provide as well as whether they meet society's needs for primary care and for service in rural and inner city areas) (Cuca et al., 1976; Hutchins, 1977).

The aim of this article is to assess the potential value in medical school admissions of the broad array of noncognitive measures, ranging from letters of recommendation and interviews to self-report questionnaires, cognitive-style tests, and objective performance devices. This effort draws not only on existing work directly concerned with admissions in such institutions, but also on the general literature dealing with assessment.

The pertinent research and theory will be cited selectively, rather than exhaustively, in view of its mass and in view of the existence of recently completed reviews, in the case of the medical admissions work (Cuca et al., 1976; Educational Environment Technical Panel, 1980; Hutchins, 1977; Otis, Albert, Quenk, Offir, Weiss, & Richardson, 1974; Pollitt, 1978).

Goals Achievable by Noncognitive Measures

In principle, noncognitive measures may contribute to achieving a variety of goals considered important in admissions: successful completion of training, adequate performance as a physician, and choice of an under-served type of practice and an under-served location (Cuca et al., 1976; Hutchins, 1977).

At least some of these same aims, of course, can also be accomplished with the help of cognitive variables. Variables of this kind, for instance, would be very pertinent to school performance, which has a large cognitive component, but would be marginally relevant to choice of location of practice, where background characteristics come into play.

Successful Completion of Training

Completion of training—broadly defined to include graduation from medical school, internship, and residency—appears to be the pre-eminent consideration in admissions and the most widely studied issue (Cuca et al., 1976). Academic work in the first two years of medical school differs fundamentally from clinical work common to the later years of medical school, internship, and residency; it is therefore important to consider the two aspects separately.

Academic work. Individual differences in performance in the first two years of medical school have been extensively investigated (Cuca et al., 1976), if only because of the ready availability of such criteria as grade-point average and dropout. It seems indisputable that noncognitive variables play a role in the academic arena, as in other spheres of life—suitable moti-

vation, for example, energizing performance and maladjustment interfering with it. Nonetheless, the ability to do academic work well appears to be largely cognitive in nature. Hence, it is not surprising that noncognitive measures have little value in predicting this criterion in medical school (Kelly, 1957).

Furthermore, insofar as noncognitive variables are implicated in academic work, they should be reflected in college course work. As a result, it is very doubtful that noncognitive measures can be identified that have an association with medical school grades not already represented in college grades, particularly those in premedical courses. For this reason, it is improbable that noncognitive measures, when combined with college grades, will improve the prediction of academic performance in medical school (Rorer, 1971).

Finally, as a practical matter, the prevalence of failure and dropout in medical school is currently so low that no pressing need exists to locate new variables that will predict these outcomes (Shulman & Elstein, 1971). Indeed, the sheer rarity of these occurrences would make them difficult to forecast.

Clinical work. The characteristics associated with performance during the last two years of medical school, during internship, and during residency have been the subject of some research but much less than that devoted to the first years of medical school; the internship and residency periods have been studied least of all (Cuca et al., 1976). This relative lack of attention reflects the absence of criteria that are easily obtained and meaningful, as well as the inherent complexity of this performance.

The demands placed on the student in these later stages of his or her training are similar to those required of the practicing physician, and the resemblance becomes progressively closer as he or she moves through the sequence from medical school to internship to residency, the major difference being that the student is inevitably more supervised than the physician.

Although it would be foolish to presume that clinical work is wholly noncognitive, a range of

noncognitive variables, from carefulness to interpersonal skills, are likely to play an important role (American Institutes of Research, 1960). In fact, the typical finding is that ability tests and grades have little or no association with clinical performance but that noncognitive measures at least have some degree of relationship (Kelly, 1966).

Performance in clinical training, as in the subsequent medical career, would seem to be an amalgam of a number of distinguishable dimensions, much more complex in nature than academic achievement or choice of type and location of practice. These dimensions would be expected to vary in the extent and character of their involvement with noncognitive variables. Consequently, it is unlikely that noncognitive variables will function monolithically with regard to clinical training, a particular measure being associated with an overall index of this performance or all of its components. It is more apt to be the case that the measures will behave differentially—some being linked with one aspect of performance, others with a second, and perhaps none with a third that is basically cognitive (Kelly, 1966).

Performance as a Physician

In view of the close similarity between the clinical stages of medical training and the subsequent practice of medicine, most of the comments made previously about training also apply to practice: (1) the question has been studied very little (Cuca et al., 1976), probably because of its difficulty; (2) noncognitive influences overshadow cognitive ones (Howell, 1966; Howell & Vincent, 1967); and (3) many different aspects of performance can be distinguished (Price, Taylor, Richards, & Jacobsen, 1964), each presumably with its own determinants.

In addition to the obvious similarities between performance in training and in practice, patent differences also exist that should have important consequences. Physicians are relatively unsupervised, as noted earlier; and the settings in which they work, even those in the same specialty and

the same type of institution, may vary considerably. Consequently, the components of performance as a physician should be even more complex, both in quantity and quality, because the uniformity produced by supervision and common settings is minimal. The role played by noncognitive variables should likewise be affected. Lack of supervision ought to increase their involvement in performance. Differences in settings, in contrast, may reduce the connection of noncognitive variables with performance, as situational characteristics assume importance and interact with them (Endler & Magnusson, 1976).

Choice of Type and Location of Practice

Choices about type of practice and its location are linked, for specialties dealing with infrequent conditions simply cannot be practiced in sparsely populated areas where the number of patients would inevitably be insufficient (Educational Environment Technical Panel, 1980). The variables involved in the choice of practice have been widely investigated (Educational Environment Technical Panel, 1980; Hutchins, 1977; Otis et al., 1974; Pollitt, 1978), but those implicated in the choice of location have been studied relatively little—most of the latter work being limited to the decision to locate in rural areas, not to the decision to locate in inner cities (Pollitt, 1978). This lack of attention to the location decision is curious, considering the clear-cut and readily available criterion as well as the important social consequences of this choice.

Choice of practice is almost necessarily determined, to a large extent, by noncognitive influences of various kinds, most particularly interests. The relationships with noncognitive variables are clearest for two specialty choices, psychiatry and surgery (Educational Environment Technical Panel, 1980; Otis et al., 1974), perhaps because these fields represent opposite extremes in their orientation towards people versus techniques (Yufit, Pollock, & Wasserman, 1969). Choice of location is also governed by noncognitive factors, especially those associated

with the person's early background (Pollitt, 1978); but it may be influenced even more by forces that have virtually nothing to do directly with the person himself or herself (Hutchins, 1977). Most notably, this choice is likely to be a joint decision of the student and his or her spouse. Insofar as such extraneous variables play a major role, the link between the student's own characteristics and the choice must be weak.

Use of Noncognitive Measures

Whatever the broad goals set for noncognitive measures in admissions, these devices are potentially usable in several ways: as predictors, as moderators, and as forecasters of admissions committee decisions. In these uses noncognitive instruments can be employed alone or in combination with cognitive measures, either supplementing or replacing the latter.

Predictors

By far the most common use of noncognitive measures in the medical admissions field, as elsewhere, is to predict criteria, and nearly all of the research follows this paradigm (Cuca et al., 1976). Interest in this application of the devices does not seem misplaced, for it may well be the most valuable way of employing them.

In gauging the value of noncognitive measures as predictors, the contribution made by cognitive instruments also needs to be considered. Simply comparing the predictability obtained with the two kinds of devices is of limited interest; more germane is examining the efficacy of combining them. Given the ultimate goal of maximizing prediction, the obvious question is whether the two in combination are more predictive than either one alone. Indeed, such combinations are often the most predictive. Whether this is the outcome in any particular instance depends on the nature of the criterion; noncognitive measures are not likely to contribute much to the prediction of cognitive criteria, nor cognitive measures to noncognitive criteria. The ex-

tent to which noncognitive measures add to predictability when combined with cognitive measures is particularly critical, for it would be difficult to justify using noncognitive instruments if they simply duplicate what is predicted by the usual kinds of cognitive devices; the latter would necessarily win in such a confrontation, if only because their use is already well established and well accepted.

Moderators

The role of noncognitive measures as moderator variables (Saunders, 1956)—variables that interact with predictors and affect their correlations with criteria—has been studied very little, especially in the medical admissions context. The prospect is naturally exciting that the relationship of ability tests with grades, for example, may differ for adjusted and for maladjusted students, making it possible to improve predictions for both groups (Hoyt & Norman, 1954; Schofield & Merwin, 1966). Nonetheless, the experience thus far, with some notable exceptions, suggests that it is difficult to identify useful moderator variables; their effects tend to be unstable, occurring in one study but not in another, and are apt to be weak even when they do appear (Stricker, 1966).

Forecasters of Admissions Committee Decisions

It has been persuasively demonstrated in recent work that the decisions of admissions committees in medical schools and in other institutions can be forecast with a substantial degree of accuracy from applicants' scores on a small number of measures, mainly admissions tests and college grades (Dawes, 1971; Milstein, Burrow, Wilkinson, & Kessen, 1976). Consequently, this set of scores, in effect, can substitute for the committees or at least can inform the groups of the relative weight that they attach to different kinds of considerations so that their operating policies can be monitored and, if need be, modified. This extremely important development

could radically alter the traditional operation of admissions committees and similar bodies, reducing the number of applicants given serious consideration and interviewed or changing the nature of the information sought and the way that it is used in making decisions (Dawes, 1971).

Although efforts to forecast committee decisions have been remarkably successful, it would obviously be desirable to maximize the accuracy of these predictions. Because committees rely on noncognitive data gleaned from letters of recommendation, interviews, and so forth (Cuca et al., 1976), appropriate noncognitive measures may contribute to the forecasts. Noncognitive devices have already been used in this research (Milstein et al., 1976), but only to a limited extent.

Kinds of Noncognitive Measures

Noncognitive measures exist in almost endless variety, nearly all of them potentially useful to one degree or another in admissions. They include a range of approaches, each represented by a large number of particular devices—some widely used and highly standardized and others rarely used, very experimental, and completely unstandardized. This discussion will be limited to approaches and measures of major interest.

Interviews

The interview is unquestionably the most universally used noncognitive measure in medical admissions, and certainly one of the most studied (Cuca et al., 1976). Paradoxically, it has long been established that the interview possesses negligible value for this purpose (Kelly, 1953). Although the interview may serve other goals of the admissions process (such as providing information to the applicant), its use as a noncognitive measure is difficult to justify on pragmatic grounds, particularly in light of the time and expense involved.

The flexibility of the interview and the opportunity it provides to observe kinds of variables not measurable in other ways are potentially im-

portant advantages. In practice, though, much of the information obtained in the interview can be secured by other means, such as application blanks and questionnaires; and the observations of the interviewer are saturated with error because of the human observer's inherent limitations (Wiggins, 1973) coupled with intentional distortion by the interviewee.

Many of these flaws might, in principle, be remedied by standardizing the interview and by using trained interviewers. The first solution, however, would destroy the interview's flexibility and increase the likelihood that it could be replaced by a questionnaire, whereas the second remedy would eliminate an important practical justification for the interview: It permits those doing the selection to meet the applicants in person.

Letters of Recommendation

Like interviews, letters of recommendation are a routine part of the admissions process in medical school (Cuca et al., 1976) and in many other institutions. These letters have not been the subject of much research. They might be expected to yield valuable information about applicants when the informants are knowledgeable. At the same time, though, letters are prone to the same errors associated with the use of human observers that affect interviews. Overt biases may also come into play when the informants, rather than being disinterested, are intent on seeing that applicants are selected. The informants are unlikely to be overly critical, in any event, because the applicants now have the legal right to see the letters. Finally, even when the letters prove to be useful, a natural question is whether the same information is not already available in some other form. For example, evaluations by premedical advisors may simply reflect the applicants' records in their undergraduate courses, which are measurable more precisely by grade-point average (Cuca et al., 1976).

Personality Inventories

Personality inventories are often used in research on the admissions process but are seldom employed in the actual selection of applicants to medical school (Cuca et al., 1976; Hutchins, 1977). Virtually the entire gamut of published instruments has been studied at one time or another. None has been employed with any frequency, but the Minnesota Multiphasic Personality Inventory (Hathaway & McKinley, 1967) and the Myers-Briggs Type Indicator (Myers, 1962) are currently used most (Hutchins, 1977).

The attraction of personality inventories is understandable, for they are an easy and inexpensive way of obtaining a broad range of potentially important information. Unfortunately, the results to date about the value of these devices in medical admissions, by and large, have not been particularly encouraging (Hutchins, 1977). This state of affairs accounts for the anomaly that inventories are widely used in research but are rarely used in practice.

It is certainly true that most of the inventories, particularly the older ones, have been given a fair trial. Nevertheless, it is premature to conclude that this entire approach has little or no merit. The technology of inventory construction has advanced dramatically in the last two decades, making it possible to devise instruments that are far superior to those of the past (Cattell, 1974; Jackson, 1970), and recent years have seen the appearance of several instruments based on this kind of up-to-date methodology. Some have already been published, such as the Personality Research Form (Jackson, 1974), Jackson Personality Inventory (Jackson, 1976), and Comrey Personality Scales (Comrey, 1970); others are nearing completion, such as the Interpersonal Style Inventory (Lorr & Youniss, 1973). These new inventories reflect the current state of the art and have considerable promise; they deserve thorough study before a final evaluation is made about the value of this mode of noncognitive measurement.

Interest Inventories

As in the case of personality inventories, the use of interest measures is fairly common in research on admissions in medical schools, but these devices are infrequently used operationally (Cuca et al., 1976; Hutchins, 1977). All of the leading instruments have been employed, but the Strong Vocational Interest Blank (SVIB; Strong, 1966) has been the most popular, at least in recent years (Hutchins, 1977). The general value of interest inventories has long been established in a variety of settings outside of medical education (Cronbach, 1970). Interest inventories would also appear to have obvious relevance in the medical context to decisions about specialty. For some inexplicable reason, though, relatively little research has dealt with this topic. An important exception is the work by Strong and Tucker (1962), which unsuccessfully attempted to devise medical specialty scales for the SVIB. This investigation raises the possibility that interest measures may be incapable of making relatively fine distinctions within an occupation.

Because the best known and most popular interest inventories of today, like the prominent personality inventories, had their inception in the 1930s when the requisite technology was comparatively primitive, the recent publication of the Jackson Vocational Interest Survey (Jackson, 1977) and the Medical Specialty Preference Inventory (Zimny, 1979) is a promising development. The former is a general purpose instrument, and the latter is designed to tap interest in common specialties. These devices, constructed with the same kind of advanced methodology used with the newer personality measures mentioned previously, have great potential.

One final observation about interest inventories is in order. Contrary to conventional wisdom, it has recently been recognized that expressions of interests elicited by direct questions (e.g., "What do you want to do?") are just as valuable as inventory measures (Baird, 1976). Direct inquiry about interests thus represents a clear and

far simpler alternative to inventory assessments of interests.

Projective Techniques

Projective techniques have never been employed as extensively as other instruments in research on medical admissions and rarely in selecting applicants; their use appears to be diminishing (Cuca et al., 1976; Hutchins, 1977). The explanation probably lies in the time and expense involved in administering and interpreting these devices as well as the spotty evidence about their value in admissions (Cuca et al., 1976). Moreover, it seems to be the case that whatever information can be obtained by these techniques can also be secured more efficiently by personality inventories and other measures (Vernon, 1964).

Before all projective techniques are written off, though, it should be recognized that certain kinds may prove to be useful in practical applications. These are devices (1) that focus on one or two variables, rather than an entire spectrum; (2) that use relatively standardized scoring procedures; and (3) that rely on surface, not depth, interpretations (Vernon, 1964). A good example is the adaptation of the Thematic Apperception Test that has been used with success in basic research in personality to assess need for achievement, need for affiliation, and other needs (Atkinson, 1958).

Biographical Measures

Biographical information, usually obtained from application blanks, is a common source of noncognitive variables in the medical school admissions process and in research on this issue (Cuca et al., 1976). Typically, single items of information are considered in isolation, often with no systematic attempt to assess their value. This state of affairs contrasts sharply with the practice in industrial settings, where it is routine to devise scales based on biographical information,

obtained from application blanks or specially devised questionnaires. These scales have been consistently found to possess substantial value in predicting job performance (Owens, 1976). Although scales of this kind have been used comparatively little in education (Baird, 1976; Freeberg, 1967) and hardly ever in other areas relevant to medical admissions (Cuca et al., 1976), on the occasions when they were employed, the results were promising (Baird, 1976; Freeberg, 1967). Anastasi, Meade, and Schneiders (1960), for example, constructed a scale that predicted achievement and adjustment in college. Even more to the point, Price, Taylor, Nelson, Lewis, Loughmiller, Mathiesen, Murray, and Maxwell (1971) developed a scale that predicted the performance of physicians.

Much of the work with biographical information has been narrowly empirical in its focus, scales being devised solely to predict a specific criterion. Consequently, it is often difficult to interpret, in general terms, the meaning of scores on the scales or to use the scales for predicting different criteria. This limitation is not inherent in biographical measures. Scales can be developed by combining homogeneous clusters of items, yielding scores that are more interpretable and perhaps of greater general use (Owens, 1976). An early example of this strategy is the Biographical Inventory for Students (Siegel, 1958).

Biographical scales, in common with personality inventories, are economical means of obtaining a large amount of information. Their unique advantage is that they capture directly the past behavior of a person, probably the best predictor of his or her future actions. Another strong point stems from the factual nature of the questions, which is less likely to lead to misinterpretation, resistance, or distortion. A potentially troublesome aspect of these devices stems from the content of the questions, for it may be inappropriate, on ethical grounds, to use information about certain kinds of characteristics, notably those over which the person manifestly has no

control, such as gender, family background, or kind of high school attended (Baird, 1976).

Cognitive Style Measures

Cognitive style measures are germane to the present discussion because of their direct implications for personality functioning. These devices are infrequently used operationally or in research on medical admissions, but they are commonly employed in psychological research generally. They exist in great variety, as do cognitive styles themselves (Messick, 1976). The most important and most studied of the cognitive styles is field independence-dependence, which is indexed by the Group Embedded Figures Test (Witkin, Oltman, Raskin, & Karp, 1971) and other devices (Witkin, Dyk, Faterson, Goodenough, & Karp, 1962). Field independence-dependence measures are associated with educational and vocational interests and even differentiate within occupations, such as clinical versus experimental psychology (Witkin, Moore, Goodenough, & Cox, 1977). However, the obvious expectation that these devices may be related to choice of medical practice was not confirmed by one major study that found that the Embedded Figures Test (Witkin et al., 1971) did not discriminate among different kinds of specialists (Goodenough, Oltman, Friedman, Moore, & Witkin, 1979).

Cognitive style measures share with ability tests the virtue of calling for maximum performance, in contrast to questionnaires and other instruments based on self-reports that attempt to assess typical performance (Cronbach, 1970). Hence, distortion is not a problem with cognitive style measures. A basic difficulty with these devices is that they may also measure ability to some extent; this extraneous influence needs to be taken into account in interpreting findings obtained with them (Cronbach, 1970).

Objective Performance Measures

Objective performance measures take many forms, probably the best known being the Lead-

erless Group Discussion (Bass, 1954). Like cognitive style measures, these devices currently play little part in admissions but are extensively used in other psychological research. It is not simple to generalize about these instruments because of their enormous heterogeneity. Like cognitive style measures, they tend to measure maximum, not typical, performance; and, unique among all instruments, they are the only ones that can faithfully mirror the behavior that must be predicted. Insofar as they successfully sample this behavior, their ability to predict it is assured. A common problem with many, though certainly not all, objective performance measures is that they are difficult to administer and to score. The results with these devices are mixed; they seem to be of some value in some cases but not in others (Cronbach, 1970; Vernon, 1964).

A number of general purpose measures exist and are potentially useful in a range of settings (Cattell & Warburton, 1967), the leading example being the Objective-Analytic Battery Test Kit (Cattell & Schuerger, 1978), which consists of 67 tests that assess 10 personality dimensions. If this measurement approach is to be exploited fully, though, it is probably necessary to tailor instruments to the specific requirements of the particular situation. An example of this strategy is the current work of Frederiksen, Ward, and Samph (1979) in devising a simulated version of the clinical interview for use in assessing the interpersonal skills of applicants to medical school.

General Problems with Noncognitive Measures

Two considerations already touched on need to be kept in mind in attempting to assess the usefulness of noncognitive measures. One concerns the criteria used to evaluate them in research on medical admissions (Hutchins, 1977). It is a truism that the results obtained in studies of this kind depend as much on the criteria as they do on the predictors. A noncognitive measure cannot be faulted for its inability to predict an inappropriate or inadequate criterion. Many

studies employ criteria that are more or less cognitive in nature (e.g., first-year grades) and thus are unlikely to be associated with a noncognitive measure. Other investigations use criteria that are complex, invalid, unreliable, or all three (e.g., ratings by supervisors) and hence not apt to be predicted by any kind of measure, cognitive or noncognitive. The criterion problem affects with equal force all instruments and all uses to which they are put.

A second consideration involves distortion in the examinees' responses arising from their desire to give a good impression. Dissembling is an ever present threat to all devices that rely on self-reports—interviews; personality and interest inventories; and, to a lesser extent, biographical measures. It occurs even in research settings under conditions of anonymity, where the incentive to distort is minimal (Edwards, 1967). In real life situations critical to the examinees, such as applying to medical school, the potential for distortion is enormous.

Left unchecked, distortion can seriously compromise self-report measures when they are used operationally. Steps can be taken to minimize dissembling in inventories and questionnaires or, at least, to identify instances in which it is prevalent. The first tactic is the better of the two, for eliminating or reducing this problem at the outset is preferable to coping with it later. Effective methods exist for constructing inventories and questionnaires resistant to distortion (Norman, 1963), but they have seldom been used. The second tactic, a common one with personality inventories, relies on special indexes to determine whether the level of distortion is excessive. When this level is reached, the record is discarded as invalid. This method, although not foolproof, works reasonably well (Jackson, 1973) but is inherently limited. It is important to know that a protocol is invalid; after this discovery has been made, however, no alternative exists but to discard the record and to remain ignorant of the examinee's characteristics.

Analogous techniques for dealing with distortion in interviews do not exist currently. It might be possible to devise methods for the identifica-

tion of dissembling, perhaps from voice records or responses to test questions; the problem of minimizing it at the outset appears to be insuperable.

References

- American Institutes of Research. *The definition of clinical competence in medicine: Performance dimensions and rationales for clinical skill areas*. Palo Alto CA: American Institutes for Research in Behavioral Sciences, 1960.
- Anastasi, A., Meade, M. J., & Schneiders, A. A. *The validation of a biographical inventory as a predictor of college success*. New York: College Entrance Examination Board, 1960.
- Atkinson, J. W. (Ed.). *Motives in fantasy, action, and society*. Princeton NJ: Van Nostrand, 1958.
- Baird, L. L. *Using self-reports to predict student performance*. New York: College Entrance Examination Board, 1976.
- Bass, B. M. The leaderless group discussion. *Psychological Bulletin*, 1954, *51*, 465-492.
- Cattell, R. B. How good is the modern questionnaire? General principles for evaluation. *Journal of Personality Assessment*, 1974, *38*, 115-129.
- Cattell, R. B., & Schuerger, J. M. *Personality theory in action—Handbook for the Objective Analytic (O-A) Test Kit*. Champaign IL: Institute for Personality and Ability Testing, 1978.
- Cattell, R. B., & Warburton, F. W. *Objective personality and motivation tests*. Champaign IL: University of Illinois Press, 1967.
- Comrey, A. L. *EITS manual for the Comrey Personality Scales*. San Diego CA: Educational and Industrial Testing Service, 1970.
- Cronbach, L. J. *Essentials of psychological testing* (3rd ed.). New York: Harper & Row, 1970.
- Cuca, J. M., Sakakeeny, L. A., & Johnson, D. G. *The medical school admissions process: A review of the literature, 1955-1976*. Washington DC: Association of American Medical Colleges, 1976.
- Dawes, R. M. A case study of graduate admissions: Application of three principles of human decision making. *American Psychologist*, 1971, *26*, 180-188.
- Educational Environment Study Panel. *Report of the Graduate Medical Education National Advisory Committee to the Secretary, Department of Health and Human Services* (Vol. 5). Hyattsville MD: U.S. Department of Health and Human Services, Office of Graduate Medical Education, 1980.
- Edwards, A. L. The social desirability variable: A review of the evidence. In I. A. Berg (Ed.), *Response*

- set in personality assessment. Chicago: Aldine, 1967.
- Endler, N.S., & Magnusson, D. (Eds.). *Interactional psychology and personality*. Washington DC: Hemisphere, 1976.
- Frederiksen, N., Ward, W. C., & Samph, T. *Development of methods for selection and evaluation in medical education*. Princeton NJ: Educational Testing Service, 1979.
- Freeberg, N. E. The biographical information blank as a predictor of student achievement: A review. *Psychological Reports*, 1967, 20, 911-925.
- Goodenough, D. R., Oltman, P. K., Friedman, F., Moore, C. A., & Witkin, H. A. Cognitive styles in the development of medical careers. *Journal of Vocational Behavior*, 1979, 14, 341-351.
- Hathaway, S. R., & McKinley, J. C. *Manual. Minnesota Multiphasic Personality Inventory* (rev. ed.). New York: Psychological Corporation, 1967.
- Howell, M. A. Personal effectiveness of physicians in a federal health organization. *Journal of Applied Psychology*, 1966, 50, 451-459.
- Howell, M. A., & Vincent, J. W. The MCAT as related to achievement tests in medicine and to supervisory evaluations of clinical physicians. *Journal of Medical Education*, 1967, 42, 1037-1044.
- Hutchins, E. B. *The selection of medical students in relation to health care needs*. Philadelphia PA: University of Pennsylvania, School of Medicine, 1977.
- Hoyt, D. P., & Norman, W. T. Adjustment and academic predictability. *Journal of Counseling Psychology*, 1954, 1, 96-99.
- Jackson, D. N. A sequential system for personality scale development. In C. D. Spielberger (Ed.), *Current topics in clinical and community psychology* (Vol. 2). New York: Academic Press, 1970.
- Jackson, D. N. Structured personality assessment. In B. B. Wolman (Ed.), *Handbook of general psychology*. Englewood Cliffs NJ: Prentice Hall, 1973.
- Jackson, D. N. *Personality Research Form manual*. Goshen NY: Research Psychologists Press, 1974.
- Jackson, D. N. *Jackson Personality Inventory manual*. Goshen NY: Research Psychologists Press, 1976.
- Jackson, D. N. *Jackson Vocational Inventory Survey manual*. Goshen NY: Research Psychologists Press, 1977.
- Kelly, E. L. An evaluation of the interview as a selective technique. In *Proceedings of the 1953 Invitational Conference on Testing Problems*. Princeton NJ: Educational Testing Service, 1953.
- Kelly, E. L. Multiple criteria of medical education and their implications for selection. *Journal of Medical Education*, 1957, 32, 10(Pt.2), 185-196.
- Kelly, E. L. Alternate criteria in medical education and their correlates. In A. Anastasi (Ed.), *Testing problems in perspective*. Washington DC: American Council on Education, 1966.
- Lorr, M., & Youniss, R. P. An inventory of interpersonal style. *Journal of Personality Assessment*, 1973, 37, 165-173.
- Messick, S. Personality consistencies in cognition and creativity. In S. Messick (Ed.), *Individuality and learning*. San Francisco: Jossey-Bass, 1976.
- Messick, S. Potential uses of noncognitive measurement in education. *Journal of Educational Psychology*, 1979, 71, 281-292.
- Milstein, R. M., Burrow, G. N., Wilkinson, L., & Kessen, W. Prediction of screening decisions in a medical school admissions process. *Journal of Medical Education*, 1976, 51, 626-633.
- Myers, I. B. *Manual. Myers-Briggs Type Indicator*. Palo Alto CA: Consulting Psychologists Press, 1962.
- Norman, W. T. Personality measurement, faking, and detection: An assessment method for use in personnel selection. *Journal of Applied Psychology*, 1963, 47, 225-241.
- Otis, G. D., Albert, M., Quenk, N., Offir, J., Weiss, J., & Richardson, C. *Medical specialty selection: A review and bibliography* (DHEW Publication No. [HRA] 75-8). Rockville MD: U.S. Department of Health, Education, and Welfare, Bureau of Health Resources Development, 1974.
- Owens, W. A. Background data. In M. D. Dunnette (Ed.), *Handbook of industrial and organizational psychology*. Chicago: Rand-McNally, 1976.
- Pollitt, A. *Social and psychological characteristics in medical specialty and geographic decisions* (DHEW Publication No. [HRA] 78-13). Hyattsville MD: U.S. Department of Health, Education, and Welfare, Manpower Supply and Utilization Branch, 1978.
- Price, P. B., Taylor, C. W., Nelson, D. E., Lewis, E. G., Loughmiller, G. C., Mathiesen, R., Murray, S. L., & Maxwell, J. G. *Measurement and predictors of physician performance: Two decades of intermittently sustained research*. Salt Lake City UT: LLR Press, 1971.
- Price, P. B., Taylor, C. W., Richards, J. M., & Jacobsen, T. L. Measurement of physician performance. *Journal of Medical Education*, 1964, 39, 203-210.
- Rorer, L. G. A circuitous route to bootstrapping. In H. B. Haley, A. G. D'Costa, & A. M. Schafer (Eds.), *Personality measurement in medical education, 1971*. Washington DC: Association of American Medical Colleges, 1971.
- Saunders, D. R. Moderator variables in prediction. *Educational and Psychological Measurement*, 1956, 16, 209-222.

- Schofield, W., & Merwin, J. C. The use of scholastic aptitude, personality, and interest test data in the selection of medical students. *Journal of Medical Education*, 1966, 41, 502-509.
- Shulman, L. S., & Elstein, A. S. Strategies for research on personality measures in medical education. In H. B. Haley, A. G. D'Costa, & A. M. Schafer (Eds.), *Personality measurement in medical education, 1971*. Washington DC: Association of American Medical Colleges, 1971.
- Siegel, L. *Manual. Biographical inventory for students*. Princeton NJ: Educational Testing Service, 1958.
- Stricker, L. J. Compulsivity as a moderator variable: A replication and extension. *Journal of Applied Psychology*, 1966, 50, 331-335.
- Strong, E. K., Jr. *Strong Vocational Interest Blanks manual*. Stanford CA: Stanford University Press, 1966.
- Strong, E. K., & Tucker, A. C. Ten-year follow-up of vocational interest scores of 1950 medical college seniors. *Journal of Applied Psychology*, 1962, 46, 81-86.
- Vernon, P. E. *Personality assessment: A critical survey*. London: Methuen, 1964.
- Wiggins, J. S. *Personality and prediction: Principles of personality assessment*. Reading MA: Addison-Wesley, 1973.
- Witkin, H. A., Dyk, R. B., Faterson, H. F., Goodenough, D. R., & Karp, S. A. *Psychological differentiation*. New York: Wiley, 1962.
- Witkin, H. A., Moore, C. A., Goodenough, D. R., & Cox, P. W. Field-dependent and field-independent cognitive styles and their educational implications. *Review of Educational Research*, 1977, 47, 1-64.
- Witkin, H. A., Oltman, P. K., Raskin, E., & Karp, S. A. *A manual for the Embedded Figures Tests*. Palo Alto CA: Consulting Psychologists Press, 1971.
- Yufit, R. J., Pollock, G. H., & Wasserman, E. Medical specialty choice and personality. I. Initial results and predictions. *Archives of General Psychiatry*, 1969, 20, 89-99.
- Zimny, G. H. *Manual for the Medical Specialty Preference Inventory*. St. Louis MO: Author, 1979.

Acknowledgments

The preparation of this article was supported in part by the Bureau of Health Manpower, Human Resources Administration, Department of Health, Education, and Welfare under Contract No. PLD 0733979. Thanks are due Norman Frederiksen and William C. Ward for critical reviews of a draft of this article.

Author's Address

Send requests for reprints or further information to Lawrence J. Stricker, Educational Testing Service, Princeton NJ 08541.