The Appraisal of

Applicants to Medical Schools

A Brief Survey of the Interview

BY JOSEPH ZUBIN*

Principal Research Scientist (Biometrics), New York State Department of Mental Hygiene

A review of the returns of the pre-Institute surveys indicates that the personal interview is the most frequently used instrument in the evaluation of the candidate for medical school—it is also the most frequent cause of griping in the student body (see Tables 4.9 and 4.11 in Chapter 4). But, despite the interview’s prevalence, much doubt exists in the minds of

---

*The author wishes to acknowledge the kind helpfulness of two of his associates at the New York State Department of Mental Hygiene: Dr. Eugene I. Burdock, Associate Research Scientist (Biometrics), and Dr. Kurt Salzinger, Research Scientist (Biometrics).
at least some of the Institute participants regarding the validity of this instrument in the selection of candidates for admission. This ambivalence toward the interview arises in the selection not only of medical students but also of other professional students and of personnel in industry; it arises in all the varied areas where the interview is customarily applied.

Paradoxically, the very people who are most eloquent about the shortcomings of the interview would never hire a secretary without an interview, nor would they ever consider engaging a research assistant or selecting a member of the medical faculty without a personal interview. Yet in all these situations the objective factors for selection could very readily be obtained either by standardized tests or from public records and papers. This paradox about the interview has been noted by many workers, including the contributors to the present symposium, and it has puzzled me for a long time—personally I would rate the challenge implicit in this paradox as the chief business of this symposium.

I am told that Dr. Aura Severinghaus of Columbia attended a similar conference about selection techniques for medical students a year or so ago in London, England. He found that he was the only one there who had anything good to say for the interview. Everyone else claimed that it was not of much value, until one New Zealander arose and said, “Gentlemen, if I wanted to get married and you had 20 willing maidens in this room and gave me all the tests in the world about them, I still would like to talk to each of them before I selected one as my wife.”

Historical background

When one reviews the history of science, it becomes quite clear that the interview, or its essence, conversation, was the first tool of scientific investigation. Long before we had yardsticks, thermometers, or clocks, man depended on the subjective responses reported by his peers or observed by himself for measuring the physical universe. This primitive, nominal measurement is still preserved in such bipolar adjectives as heavy-light, warm-cold, long-short, and so on, and their counterparts in other languages. Only the provision of criteria external to the subjective estimates of human beings has made it possible to objectify the measuring of distance, heat, time and so on.

In psychology, too, the measurement of human behavior was initially self-referred, and this nominal stage in psychological measurement reached its climax with the checklist of descriptive adjectives, such as that of Cattell. Even today with the availability of a variety of objective tools the bright student in a class is the one most like the instructor, while the dull fellow can be identified by his failure to appreciate the professor’s point of view. Although this system of self-reference works pretty well most of the time, it occasionally results in gross incongruities. Thus Evariste Galois, who invented the mathematical theory of groups, was declared to be hopelessly incompetent by his professors and was refused admission to the Sorbonne.
The concept of personality originally included all the traits and characteristics that had anything to do with the evaluation of human behavior, including intelligence, interests, and aptitudes. But then Binet introduced the intelligence test, which Stern subsequently capped with that ubiquitous index of intellectual efficiency, the IQ. Subsequently E. K. Strong devised a measure of vocational interests by weighting the contrasting preferences of different occupational groups on an interest questionnaire. About the same time, R. S. Woodworth brought out the first personality inventory for the assessment of personal adjustment; Carl Seashore developed a measure of musical aptitude, and other special aptitudes were found to be accessible to independent assessment.

Through these developments personality gradually lost its "intelligence," then its "interests," and finally even its "aptitudes." Nevertheless, the concept of personality could not be dispensed with, because there were too many individual differences left in behavior after the measurable components had been accounted for. These could only be evaluated by means of conversation or interview. That is why the interview has persisted despite the fact that it has suffered from a social lag in comparison with the other instruments we use for the evaluation of behavior. For, while standardized tests, psychophysical techniques, and psychological methods have become more and more sophisticated, the interview has remained a vague, global assessment device—a catch basin for the residue of behavior that filters through the network of objective measures.

Components of the interview process

In order to see how this bear cub, the interview, may be licked into respectable scientific shape, one must first analyze the interview process into its components and learn how to manipulate the variables involved so as to elicit the kind of information desired. The four essential features of the interview are:

1. It is a face-to-face situation
2. It has a specific purpose
3. It uses interpersonal conversation as a medium of communication
4. The relationship between the two interlocutors is not reciprocal (i.e. the roles of interviewer and interviewee are not interchangeable)

The elimination of any one of these features necessarily reduces the interview to a degenerate form (mathematical not human degeneration). For example, eliminating the face-to-face situation, without disturbing any of the other elements, withdraws certain intrinsic gestural cues. The effects of cue reduction are evident in the telephone interview or at the confessional. Unilateral modification of the face-to-face situation, as when one of the two participants is unaware that he is being observed, reduces the interview to an observational procedure in which both verbal and nonverbal behavior are recorded by an observer. The result of removing conversational give-and-take with the retention of the face-to-face situation is seen in the lecture or in the psychoanalytic session. The removal of conversation and substitution of written responses reduces the interview to
a supervised test session. If specific purpose is absent, the interview is reduced to conversation or gossip or-bull session.

No systematic variation of the essential features of the interview has thus far been attempted. However, for our purposes today, it might be well to limit ourselves to one important feature of the interview, namely its specific purpose. The purpose of the selection interview ought not to be determination either of the intellectual capacities of the candidate or of any other aspects of his behavior that can be more accurately gauged by standardized tests, official records, and so on. We should reserve the interview for those things which are not attainable by any other method. And, what are these? I would like to suggest that the purpose of the interview is to gather information about the candidate's motives, feelings, attitudes, and integrity insofar as they determine his interests in medicine and his ability to deal with people.

These are factors for which no tests exist today, although perhaps some day we may be able to develop tests for them, if we obtain suitable criteria. Whether or not these factors are important for eventual success in medicine is, of course, still debatable, but it seems that the majority of the members of this Association would like to have information on these factors and that is why I draw attention to them. As for any connection between these factors and the eventual outcome of medical training, we shall have to isolate more adequate criteria of outcome before we can expect to devise tests that will be criterion-relevant. Moreover, although it is quite likely that the traits and characteristics we have mentioned will eventually be important for success in medicine, should they turn out to be unimportant, much of our present point of view about professional selection will prove to have been wrong.

If we agree that the purpose of the interview is to focus on the personality characteristics which are important and at the present time cannot be measured in any other way, the next question is: How does the interviewer elicit such information? The best answer that we can give at the present time is that the interviewer operates with the tactics and strategy of ordinary conversation. These tactics are overlearned behaviors that are reinforced repeatedly from the time when we are participating in social conversation. It is therefore very difficult to disengage oneself from the habits of ordinary conversation for the purpose of conducting an interview. That is why it is necessary to train interviewers. Only those who have explicit understanding of the techniques involved in conversation can properly manage the interview situation. Some of these techniques can be cataloged, although not all of them.

For example, if you were to observe yourself or others engaged in conversation, you would note that attitudes, feelings, and motivations were often evoked by the method of amplified agreement. Whenever the interviewer wanted to find out whether the point of view expressed by the interlocutor was really representative of the general view that he had, he would extend the statement by amplifying it. The interviewee would then either correct the amplification, ignore it, or agree with it. In this way, the limitations
of his attitudes and the tenacity with which he adhered to them would be revealed to the observant interviewer.

Other techniques are disagreement, surprise, change of tempo, and so forth, tactics we are all capable of employing and that serve us well in conversational situations. One might almost go so far as to say that the ability to adopt a variety of attitudes toward the interviewee is essential for the good interviewer.

Some pitfalls of the interview

Errors of inference often arise in the course of manipulating one’s own attitudes to elicit those of an interlocutor. Interviewers may arrive at invalid conclusions because of (1) erroneous information and (2) emotional prejudices. These two kinds of errors can be moderated only by experience and training.

A third type of error is a psychological effect of the interview process itself. Thorndike pointed out long ago that when ratings of specific traits are made by means of interviews the correlations between traits are quite high, even where there is no a priori reason for assuming any relation between the traits. For example, when ratings on such traits as friendliness, studiousness, and dancing ability are correlated, the correlations are sometimes surprisingly high, even though no relation between these three traits is to be expected. This so-called halo effect is usually attributed to a generalized attitude of the rater towards the subject which modifies all of his judgments.

A fourth type of bias that has been discovered in the interview is a tendency toward stereotypy. Walter Lippmann introduced this term to describe the public’s inclination to think in, and to respond to, stereotypes during political elections. Allport, who called this tendency the “labeling error,” pointed out that it represents a psychological paralysis of judgment through oversimplification. For example, the assumption that students who have majored in the biological sciences make the best physicians may represent such a labeling error.

A fifth error arises from the predominance of ratings near the center of the judgment scale because of the judge’s attempt to avoid extreme ratings. A rater who is not familiar with the statistical theory of distributions may not realize that if he places everybody at the mean, or at the center, his work is undone. For he has provided no information for differentiating the population.

A sixth error is the generosity effect which tends to make the rater or interviewer give the subject the benefit of the doubt. The generosity effect, when coupled with certain biases on the part of the interviewer, can often lead him astray. For example, clinical psychologists and psychiatrists often lean over backwards to give a candidate the benefit of the doubt in situations where businessmen would not hesitate to reject a candidate. Clinical experience may make us feel so much more tolerant of deviant behavior that we lose our perspective. Such relaxed standards may result in minimizing a candidate’s dishonesty or other lack of integrity.
A seventh type of error is the occasional suggestive nature of interviewer questions which steer the response of the interviewee. This may happen when the interviewer has a strong bias towards a given hypothesis. The interviewer may then inadvertently suggest answers to the interviewee.

Finally, errors may arise as a result of improper weighting of the different elements of the interview when a composite picture is later assembled by the interviewer. Sometimes in retrospect a certain trait will be recalled that turned out to be crucial, although it was deemed of no significance at the time of the interview.

That these sources of error are potent in influencing the conclusions based on interviews is strictly evidenced by the results obtained in a special study of one of our medical schools. All the interviewer ratings of personality, but not the ratings on evaluation of past scholastic achievement, were negatively correlated with actual success in training. One telltale indication of why such interviews are ineffectual is the naïve assumption of many interviewers that the first impression made during the interview is important. Sixty-two per cent of the admissions committee members who responded to the pre-Institute questionnaire believed this to be true (see Table 5.1). Moreover, 60 per cent of the members rated their confidence in the accuracy of their initial judgment as fairly high. Since the accuracy of first impressions was disproved several decades ago, such beliefs belong to folklore rather than to science. Unless we train interviewers to be aware of the above sources of error, negative correlations between interview and criterion are likely to keep occurring indefinitely.

**Table 5.1**

ADMISSION COMMITTEE MEMBERS' OPINIONS ON THE FIRST IMPRESSION OF AN APPLICANT IN AN INTERVIEW

<table>
<thead>
<tr>
<th>Importance of first impression</th>
<th>None</th>
<th>Little</th>
<th>Some</th>
<th>Much</th>
<th>No response</th>
<th>% rating confidence in their first impression</th>
<th>% rating importance of first impression</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>*</td>
<td>0</td>
<td>5</td>
<td>61</td>
</tr>
<tr>
<td>Little</td>
<td>3</td>
<td>17</td>
<td>6</td>
<td>3</td>
<td>1</td>
<td>20</td>
<td>51</td>
</tr>
<tr>
<td>Some</td>
<td>1</td>
<td>9</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>91</td>
<td>51</td>
</tr>
<tr>
<td>Much</td>
<td>6</td>
<td>*</td>
<td>6</td>
<td>*</td>
<td>3</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td>No response</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>3</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Total (N = 834)</td>
<td>6</td>
<td>28</td>
<td>47</td>
<td>13</td>
<td>6</td>
<td>105</td>
<td></td>
</tr>
</tbody>
</table>

*Less than 1 per cent response.

**Some interview techniques**

So much for the interview in relation to our usual conversational techniques. Let me take a few moments to draw attention to some other approaches to the interview.

First, there is the practice of recording the interview for the purpose of making a content analysis in accordance with certain specified categories. Content analysis is a well-developed technique. If we could obtain tape recordings of admission interviews, we might analyze their content and play them back to the interviewers pointing out what areas were covered, what areas were missed, and what sources of error were in evidence. Correlations could then be run between the results of content analysis and
the interviewer's evaluation. I know of only one medical school that records interviews. If such records could be made available for content analysis, the whole selection process would be the beneficiary.

An experimental approach to the interview is afforded by the direct and conscious manipulation of certain elements found in all conversation. Usually we apply these techniques in an unconscious and natural way. Methods have recently been developed for varying them systematically. Thus, in one experimental form of the interview, the interviewer varies his "input" according to a predetermined schedule. At fixed intervals during the interview the interviewer will remain silent, or interrupt the interviewee, or interfere in some other way with the smooth flow of conversation. The response of the interviewee to these conditions reveals many of his personality characteristics.

Chapple has shown such techniques to be efficacious in the selection of salesmen of ladies shoes. Saslow and Matarazzo, basing their work upon that of Chapple, have demonstrated that the temporal aspects of the interview revealed characteristics of the interviewee. These investigators have established the reliability of their findings in well-controlled and replicated studies.

Another experimental approach to the interview has utilized content-free speech. The procedure is to take the recording and filter out from the interviewee's voice the audible frequencies below 450 cycles per second. This is done by means of a high-pass electronic filter. The resulting speech record still yields indications of pitch, rate, and loudness—the voice qualities that are often said to be the best indicators of affect or emotion. Starkweather, Kauffman, and Soskin have been able to demonstrate correlations between speech qualities and such traits as pleasantness and aggressiveness.

A more recent innovation is the focused interview, intended to elicit specific information about a given topic. For example, in our own recent work on prognosis of outcome of schizophrenia, we have been developing a technique for determining the plasticity of the patient's affective communication. As some of you no doubt know, flatness of affect is an indicator that augurs ill for outcome. Most patients with flattened affect remain in the hospital for the rest of their lives.

In order to obtain a measure of flatness, we record a type of interview like the one described here. During the first ten minutes of interview the patient is encouraged to discuss any area of his experience, while the inter-

---


viewer maintains a noncommittal, detached attitude. A count is taken of the number of utterances with affective words such as "like," "dislike," "love," "hate," or any that refer to a subjective feeling. During the second ten-minute period every time the patient utters a reference to some subjective feeling or affect, the examiner encourages or reinforces the utterance by saying, "Uh-huh," or "Yes." Invariably, the number of affective utterances rises precipitously. In the final ten minutes the experimenter reverts to his initial neutral tactics, so as to discourage continued expression of affect. The variation from patient to patient in ability to emit affect is most striking. We think we have obtained a measure that is at once objective and suitable for prognostic purposes.

There are implications here for the selection interview. One might, for example, focus on some aspect of the candidate's attitude or experience or inclinations. If every time the candidate made a reference to a motivational factor or to an attitude toward medicine the examiner were to say "Uh-huh" or "Yes," the record obtained in this fashion might later be quite revealing about the variety of motivation the candidate possessed, the quality of these motivations, and their relevance to later success.

A final word

The purpose of this brief survey of the development of interview techniques has been to set the stage for the three papers that follow and to suggest some approaches to interviewing that may reduce the hazards involved in using this procedure in selection. It is hoped this symposium contributes some ideas that medical schools can utilize in handling the paradoxical attitudes toward interviewing their applicants for admission.