

Chapter 3

THE ASSESSMENT INTERVIEW

Probably the single most important means of data collection during psychological evaluation is the assessment interview. Without interview data, most psychological tests are meaningless. The interview also provides potentially valuable information that may be otherwise unobtainable, such as behavioral observations, idiosyncratic features of the client, and the person's reaction to his or her current life situation. In addition, interviews are the primary means for developing rapport and can serve as a check against the meaning and validity of test results.

Sometimes an interview is mistakenly thought to be simply a conversation. In fact, the interview and conversation differ in many ways. An interview typically has a clear sequence and is organized around specific, relevant themes because it is meant to achieve defined goals. Unlike a normal conversation, the assessment interview may even require the interviewer and interviewee to discuss unpleasant facts and feelings. Its general objectives are to gather information that cannot easily be obtained through other means, establish a relationship that is conducive to obtaining the information, develop greater understanding in both the interviewer and interviewee regarding problem behavior, and provide direction and support in helping the interviewee deal with problem behaviors. The interviewer must not only direct and control the interaction to achieve specific goals, but also have knowledge about the areas to be covered in the interview.

A basic dimension of an interview is its degree of structure. Some interviews allow the participants to freely drift from one area to the next, whereas others are highly directive and goal oriented, often using structured ratings and checklists. The more unstructured formats offer flexibility, possibly high rapport, the ability to assess how clients organize their responses, and the potential to explore unique details of a client's history. Unstructured interviews, however, have received frequent criticism, resulting in widespread distrust of their reliability and validity. As a result, highly structured and semistructured interviews have been developed that provide sound psychometric qualities, the potential for use in research, and the capacity to be administered by less trained personnel.

Regardless of the degree of structure, any interview needs to accomplish specific goals, such as assessing the client's strengths, level of adjustment, the nature and history of the problem, diagnosis, and relevant personal and family history. Techniques for accomplishing these goals vary from one interviewer to the next. Most practitioners use at least some structured aids, such as intake forms that provide identifying data and basic elements of history. Obtaining information through direct questions on intake forms frees the clinician to investigate other aspects of the client in a more flexible, open-ended manner. Clinicians might also use a checklist to help ensure that they have

covered all relevant areas. Other clinicians continue the structured format throughout most of the interview by using one of the formally developed structured interviews, such as the Schedule for Affective Disorders and Schizophrenia (SADS) or Structured Clinical Interview for the *DSM-IV* (SCID).

HISTORY AND DEVELOPMENT

Early Developments

The earliest form of obtaining information from clients was through clinical interviewing. At first, these interviews were modeled after question-and-answer medical formats, but later, the influence of psychoanalytic theories resulted in a more open-ended, free-flowing style. Parallel to the appearance of the psychoanalytically oriented interview was the development of the more structured and goal-oriented mental status examination originally formulated by Adolf Meyer in 1902. The mental status examination assessed relevant areas of a client's current functioning, such as general appearance, behavior, thought processes, thought content, memory, attention, speech, insight, and judgment. Professionals also expressed early interest in the relationship between biographical data and the prediction of occupational success or prognosis for specific disorders.

Regardless of the style used, the interviews all had these common objectives: to obtain a psychological portrait of the person, to conceptualize what is causing the person's current difficulties, to make a diagnosis, and to formulate a treatment plan. The difficulty with unstructured interviews is that they were (and still are) considered to have questionable reliability, validity, and cost-effectiveness. The first standardized psychological tests were developed to overcome these limitations. Tests could be subjected to rigorous psychometric evaluation and were more economical because they required less face-to-face contact with the person(s) being evaluated.

Developments during the 1940s and 1950s

During the 1940s and 1950s, researchers and clinicians began conceptualizing and investigating the following critical dimensions of interviews:

1. Content versus process.
2. Goal orientation (problem solving) versus expressive elements.
3. Degree of directiveness.
4. Amount of structure.
5. The relative amount of activity expressed by the participants.

These issues have been the focus of numerous research studies. A representative and frequently cited study on interviewer style was reported by W. Snyder (1945), who found that a nondirective approach was most likely to create favorable changes and self-exploration in clients. In contrast, a directive style using persuasion, interpretation, and

interviewer judgments typically resulted in clients being defensive and resistant to expressing difficulties. Strupp (1958) investigated the experience-inexperience dimension and found, among other things, that experienced interviewers expressed more warmth, a greater level of activity, and a greater number of interpretations. Level of empathy did not alter, regardless of the interviewer's degree of experience. Further, representative studies include Porter's (1950) in-depth evaluation of the effects of different types of responses (evaluative, probing, reassuring) and R. Wagner's (1949) early review, which questioned the reliability and validity of employment interviews.

Developments during the 1960s

A considerable amount of research in the 1960s was stimulated by C. Rogers (1961), who emphasized understanding the proper interpersonal ingredients necessary for an optimal therapeutic relationship (warmth, positive regard, genuineness). Elaborating on Roger's ideas, Truax and Carkhuff (1967) developed a five-point scale to measure interviewer understanding of the client. This scale was used for research on interviewing, therapist training, and as support for a client-centered theoretical orientation. Additional research efforts were also directed toward listing and elaborating on different categories of interactions such as clarification, summarizing, and confrontation.

Other investigators conceptualized interviewing as an interactive system in which the participants simultaneously influenced each other (Matarazzo, 1965; Watzlawick, Beavin, & Jackson, 1966). This emphasis on an interactive, self-maintaining system became the core for most early and later formulations of family therapy. The 1960s also saw the development and formalization of behavioral assessment, primarily in the form of goal-directed interviews that focused on understanding current and past reinforcers as well as on establishing workable target behaviors. Proponents of behavioral assessment also developed formal rating instruments and self-reports for areas such as depression, assertiveness, and fear.

Some attempts were made at integrating different schools of thought into a coherent picture, such as Beier's (1966) conceptualization of unconscious processes being expressed through nonverbal behaviors that could then be subject to covert social reinforcement. However, the 1960s (and part of the 1970s) were mostly characterized by a splintering into different schools of conflicting and competing ideologies. For example, client-centered approaches emphasized the importance of staying with the client's self-exploration; behavioral interviews emphasized antecedents and consequences of behavior; and family therapy focused on interactive group processes. Parallel progress was made within each of these different schools and within different disciplines, but little effort was devoted to cross-fertilization and/or integration.

Throughout the 1950s and 1960s, child assessment was conducted primarily through interviews with parents. Direct interviews with the child were considered to be for therapeutic purposes rather than for assessment. Differential diagnosis was unusual; almost all children referred to psychiatric clinics were either undiagnosed or diagnosed as "adjustment reactions" (Rosen, Bahn, & Kramer, 1964). Early research by Lapouse and Monk (1958, 1964) using structured interviews, indicated that mothers were more likely to report overt behaviors that are bothersome to adults (thumb-sucking, temper tantrums), but children were more likely to reveal covert difficulties (fears, nightmares).

Somewhat later, P. Graham and Rutter (1968), using structured interviews of children (rather than a parent), found interrater agreement was high for global psychiatric impairment (.84); moderate for attentional deficit, motor behavior, and social relations (.61 to .64); and low for more covert difficulties such as depression, fears, and anxiety (.30).

Developments during the 1970s

Assessment with adults and children during the 1970s saw a further elaboration and development of the trends of the 1960s, as well as increased emphasis on structured interviews. The interest in structured interviews was fueled largely by criticisms about the poor reliability of psychiatric diagnosis. A typical structured interview would be completed by the interviewer either during or directly after the interview, and the data would be transformed into such scales as organicity, disorganization, or depression-anxiety.

Initial success with adult structured interviews (e.g., Present State Examination, Rernard Diagnostic Interview) encouraged thinking regarding the further development of child-structured interviews both for global ratings and specific content areas. Child assessment became concerned not only with information derived from parents, but also with the child's own experience. There was a trend toward direct questioning of the child, greater emphasis on differential diagnosis, and the development of parallel versions of structured interviews for both the parent(s) and child.

Behavioral strategies of interviewing for both children and adults not only emphasized the interviewee's unique situation, but also provided a general listing of relevant areas for consideration. Kanfer and Grimm (1977) outlined the areas an interviewer should assess as:

1. Behavioral deficiencies.
2. Behavioral excesses.
3. Inappropriate environmental stimulus control.
4. Inappropriate self-generated stimulus.
5. Problem reinforcement contingencies.

In a similar categorization, Lazarus (1973) developed his BASIC-ID model, which describes a complete assessment as involving behaviors (B), affect (A), sensation (S), imagery (I), cognition (C), interpersonal relations (I), and need for pharmacological intervention/drugs (D).

Additional themes in the 1970s included interest in biographical data, online computer technology, and the training of interviewer skills. Specifically, efforts were made to integrate biographical data for predicting future behavior (suicide, dangerousness, prognosis for schizophrenia) and for inferring current traits. J. W. Johnson and Williams (1977) were instrumental in developing some of the earliest online computer technology to collect biographical data and to integrate it with test results. Although training programs were devised for interviewers, a central debate was whether interview skills could actually be significantly learned or improved (Wiens, 1976).

Whereas most reviews of the literature in the 1970s emphasized the advantages of a comprehensive structured format, family therapists were dealing with group processes

in which formal interview structure was typically deemphasized. Because most family therapists were observing fluid interactional processes, they needed to develop a vocabulary different from that used in traditional psychiatric diagnosis. In fact, *DSM* categories were usually considered irrelevant because they described static characteristics of individuals rather than ongoing group processes. Few, if any, structured formats were available to assess family relationships.

Developments during the 1980s

Many of the trends, concepts, and instruments developed in the 1960s and 1970s were further refined and adapted for the 1980s. One important effort was the adaptation of many instruments to the *DSM-III* (1980) and *DSM-III-R* (1987). In addition, the increased delineation of childhood disorders required greater knowledge related to differential diagnosis and greater demand for structured interviews as adjuncts to assessment. Many of the efforts were consistent with the use of specific diagnostic criteria along with a demand for efficiency, cost-effectiveness, and accountability. Despite concerns regarding computer-based interpretations (Groth-Marnat & Schumaker, 1989), some of these functions were beginning to be performed by specific computer programs. Because interviews were becoming increasingly structured, with the inclusion of scales and specific diagnostic strategies, the distinction between tests and interviews was becoming less clear. In some contexts, aspects of interviewing were even replaced with computer-requested and computer-integrated information and combined with simple programs to aid in diagnosis, such as DIANO III (Spitzer, Endicott, & Cohen, 1974) and CATEGO (Wing, Cooper, & Sartorius, 1974). During the mid- and late 1980s, most clinicians, particularly those working in large institutions, used a combination of structured interviews and open-ended unstructured approaches. Some research focused on the importance of the initial interview regarding clinical decision making and later therapeutic outcome (Hoge, Andrews, Robinson, & Hollett, 1988; Turk & Salovey, 1985). There was also a greater appreciation and integration of the work from different disciplines and from differing theoretical persuasions (Hersen, 1988). Finally, greater emphasis was placed on the impact and implications of culture and gender on the assessment process (L. Brown, 1990).

The 1990s and Beyond

Two of the defining features of psychology in the 1990s were managed health care and the controversy over the validity of repressed memories. Both of these issues had significant implications for interviewing. Managed health care emphasized the cost-effectiveness of providing health services; and for interviewing, this means developing the required information in the least amount of time. This may mean streamlining interviews by maximizing computer-derived information or paper-pencil forms. This brings up the larger issue of the extent to which practitioners need to spend face-to-face time with the client versus deriving information through other means. The development of single-session therapy (Hoyt, 1994) illustrates the potential brevity of information that might be required before making therapeutic interventions. There was also recognition that precise patient-treatment matching can optimize the treatment

and potentially the cost-effectiveness of psychosocial interventions (Antony & Barlow, 2002; Beutler & Clarkin, 1990; Beutler, Clarkin, & Bongar, 2000).

The controversy over repressed memories has forced interviewers to clarify the extent to which the information they derive from clients represents literal as opposed to narrative truth. Research has consistently indicated that client self-reports are reconstructions of events (Henry, Moffitt, Caspi, Langley, & Silva, 1994; Loftus, 1993) and are likely to be particularly questionable for retrospective reports of psychosocial variables (Henry et al., 1994). The even greater challenge to interviewers is to ensure that their interviewing style and method of questioning are not distorting the information derived from clients. This issue becomes intensely highlighted during interviews to investigate the possibility of childhood sexual abuse (see guidelines in S. White & Edelstein, 1991).

Further, continuing themes in the 1990s were the impact of gender and cultural issues and the further development of structured interviews. In some cases, the preceding issues have produced tension. For example, the greater demands for brief focused interventions contradict the emphasis of structured interviews on detailed and often time-consuming procedures. In addition, there has been greater clinical and political importance attached to detecting and treating childhood abuse; yet research and media coverage of recovered memories have suggested that some, if not many, of these memories are of questionable validity. The themes related to cost-effectiveness, patient-treatment matching, recovered memories, use of structured interviews, and cultural and gender issues are far from resolved and will continue to be important themes during this decade.

ISSUES RELATED TO RELIABILITY AND VALIDITY

Although the interview is not a standardized test, it is a means of collecting data and, as such, can and should be subjected to some of the same types of psychometric considerations as a formal test. This is important because interviews might introduce numerous sources of bias, particularly if the interviews are relatively unstructured. Reliability of interviewers is usually discussed in relation to interrater (interviewer) agreement. R. Wagner's (1949) early review of the literature found tremendous variation, ranging from .23 to .97 (*Mdn* = .57) for ratings of personal traits and -.20 to .85 (*Mdn* = .53) for ratings of overall ability. Later reviews have generally found similar variations in interrater agreement (Arvey & Campion, 1982; L. Ulrich & Trumbo, 1965). The problem then becomes how to determine which ratings to trust and which to view with skepticism. Of particular relevance is why some interviewers focus on different areas and have different biases. A consistent finding is that, when interviewers were given narrow areas to assess and were trained in interviewer strategies, interrater reliability increased (Dougherty, Ebert, & Callender, 1986; Zedeck, Tziner, & Middlestadt, 1983). The consensus was that highly structured interviews were more reliable (Huffcutt & Arthur, 1994; McDaniel, Whetzel, Schmidt, & Maurer, 1994). However, increased structure undermines one of the greatest strengths of interviews—their flexibility. In many situations, a free-form, open-ended approach may be the only way to obtain some types of information.

Research on interview validity has typically focused on sources of interviewer bias. For example, halo effects result from the tendency of an interviewer to develop a general impression of a person and then infer other seemingly related characteristics. For example, clients who are considered to express warmth may be seen as more competent or mentally healthy than they actually are. This clustering of characteristics may be incorrect, thereby producing distortions and exaggerations. Similarly, first impressions have been found to bias later judgments (W. Cooper, 1981). Confirmatory bias might occur when an interviewer makes an inference about a client and then directs the interview to elicit information that confirms the original inference. For example, a psychoanalytically-oriented interviewer might direct questions related to early childhood traumas, possibly incorrectly confirming traditional psychoanalytic explanations of current adult behaviors. Similar to halo effects is the finding that one specific outstanding characteristic (educational level, physical appearance, etc.) can lead an interviewer to judge other characteristics that he or she incorrectly believes are related to the outstanding one. For example, physical attractiveness has been found to create interviewer bias in job applicants (Gilmore, Beehr, & Love, 1986). In a clinical context, physical attractiveness may result in practitioners' either deemphasizing pathology or, on occasion, exaggerating pathology because of discomfort the interviewer may feel over his or her feelings of attraction (L. Brown, 1990). Interviewers also may focus incorrectly on explanations of behavior that emphasize traits rather than situational determinants (Ross, 1977). This error is particularly likely when the interpretation of interview data relies heavily on psychological tests, because tests, by their nature, conceptualize and emphasize static characteristics of the person rather than ongoing interactional processes.

In addition to the interviewer's perceptual and interactional biases, the interviewees may distort their responses. For example, they may present an overly favorable view of themselves, particularly if they are relatively naive regarding their motivations. Distortions are most likely found in sensitive areas such as sexual behavior. Some specific areas of distortions are represented by the finding that victims of automobile accidents typically exaggerated the amount of time they lost from work, 40% of respondents provided overestimates of their contributions to charity, and 17% of respondents reported their ages incorrectly (R. Kahn & Cannell, 1961). More extreme cases of falsification occur with outright (conscious) lies, delusions, confabulations, and lies by pathological (compulsive) liars that they partially believe themselves (Kerns, 1986). Inaccuracies based on retrospective accounts have been found to most likely occur related to psychosocial information (e.g., family conflict, onset of psychiatric symptoms) compared with variables such as change of residence, reading skill, height, and weight (B. Henry et al., 1994).

Reviews of interview validity, in which interviewer ratings were compared with outside criterion measures, have, like reliability measures, shown tremendous variability ranging from -0.05 to $+0.75$ (Arvey & Campion, 1982; Henry et al., 1994; Huffcutt & Arthur, 1994; J. Hunter & Hunter, 1984; L. Ulrich & Trumbo, 1965). One clear finding is that validity increases as the structure of the interview format increases (Huffcutt & Arthur, 1994; Marchese & Muchinsky, 1993). For example, a meta-analysis by Wiesner and Cronshaw (1988) found that unstructured interviews had validity coefficients of .20,

structuring the interview increased the validity to .63, and structured interviews by a panel using consensus ratings increased validity coefficients to a quite respectable .64. However, the validity seems to vary according to the type of variable that is being assessed. Situational employment interviews (asking the interviewee what he or she would do in a particular situation) had higher validities (.50) than interviews used to assess past job-related behavior (.39) or rate psychological qualities such as dependability (.29; McDaniel et al., 1994). It has also been found that interview accuracy increases more when interviewees are held accountable for the process they went through when coming to their decisions, compared to being held accountable for the accuracy of their predictions (procedural versus outcome accountability; Brtek & Motowidlo, 2002).

The previous brief review indicates that adding structure to interviews and paying close attention to the procedure by which decisions are made typically results in higher levels of validity. It also means that information derived from unstructured interviews should be treated cautiously and treated as tentative hypotheses that need to be supported by other means. Interviewers should also continually question the extent to which their particular style, attitudes, and expectations might be compromising interview validity. Given the difficulties related to unstructured formats, a variety of formal structured clinical interviews has been developed. Additional information on the reliability and validity of the most frequently used structured clinical interviews is provided in the last section of this chapter.

ASSETS AND LIMITATIONS

Both structured and unstructured interviews allow clinicians to place test results in a wider, more meaningful context. In addition, biographical information from interviews can be used to help predict future behaviors; what a person has done in the past is an excellent guide to what he or she is likely to continue doing in the future. Factors for predicting suicide risk, success in certain occupations, and prognosis for certain disorders can usually be most effectively accomplished by attending to biographical data rather than test scores. Because tests are almost always structured or “closed” situations, the unstructured or semistructured interview is typically the only time during the assessment process when the clinician can observe the client in an open, ambiguous situation. Observations can be made regarding how persons organize their responses, and inferences can be derived from subtle, nonverbal cues. These inferences can be followed up with further, more detailed questioning. This flexibility inherent in unstructured and semistructured interviews is frequently their strongest advantage over standardized tests. The focus during unstructured interviews is almost exclusively on the individual rather than on how that individual does or does not compare with a larger normative comparison group. Some types of information can be obtained only through this flexible, person-centered approach, which allows the interviewer to pay attention to idiosyncratic factors. In crisis situations when relatively rapid decisions need to be made, it can be impractical to take the time required to administer and interpret tests, leaving interviews and rapid screening devices as the only means of assessment. Finally, interviews allow clinicians to establish rapport and encourage client self-exploration. Rarely do

clients reveal themselves nor do they perform optimally on tests unless they first sense trust, openness, and a feeling of being understood.

The greatest difficulty with unstructured interviews is interviewer bias from perceptual and interactional processes such as the halo effect, confirmatory bias, and the primacy effect. This bias typically results in considerable variability for both reliability and validity as well as in difficulty comparing one subject with the next. One of the main reasons for diagnostic disagreement is variations in the information obtained (information variance) and variations in the criteria (criterion variance) used to conclude the presence or absence of a condition. In more concrete terms, this means that different practitioners develop and ask a wide variety of questions and apply standards for the presence of a condition, such as depression, in an inconsistent fashion. A further difficulty is the high cost of using trained interviewers for large-scale epidemiological studies.

Structured interviews have many distinct advantages over unstructured approaches. Because structured interviews have more psychometric precision, the results enable comparability between one case or population and the next. The standardized presentation allows for the development of reliable ratings, reduces information variance, and uses consistent diagnostic criteria (R. Rogers, 1995; Summerfeldt & Antony, 2002). In addition, the comprehensiveness of many structured interviews reduces the likelihood of missing a diagnosis or set of relevant symptomology. Partially because of these advantages, structured clinical interviews have progressed from being used primarily for research to use in a number of clinical settings. At issue, however, is the time required for structured interviews. The more recently developed computer-assisted programs offer a potential method of countering this difficulty (Epstein & Klinkenberg, 2001). In addition, instruments such as the Diagnostic Interview Schedule and Diagnostic Interview for Children and Adolescents have been designed for administration by lay interviewers, thereby reducing the time required by professionals.

Although structured interviews generally have higher psychometric properties than unstructured formats, they tend to overlook the idiosyncrasies and richness of the person. In many cases, these unique aspects may go undetected and yet may make a significant difference in interpreting test scores or making treatment recommendations. Although still somewhat controversial (Helzer & Robins, 1988), another criticism of many clinicians and researchers is that a highly structured approach may not create enough rapport for the client to feel sufficiently comfortable about revealing highly personal information. This is truer for the highly structured interviews, such as the Diagnostic Interview Schedule, than for a semistructured instrument, such as the Schedule for Affective Disorders and Schizophrenia, which includes an initial, relatively unstructured component. However, M. Rosenthal (1989) has noted that rapport with structured instruments can be enhanced through carefully educating the client as to the importance and procedures of these more structured approaches.

Although many of the structured interviews have demonstrated adequate reliability, studies relating to validity have primarily focused on the general level of impairment or simple discriminations between psychiatric and nonpsychiatric populations. There has been considerable controversy over what exactly is an acceptable outside criterion measure regarding the "true" diagnosis. In-depth studies of construct validity or incremental validity have yet to be performed. Furthermore, far more work needs to be done

on the treatment utility of structured interviews in areas such as prognosis, selection of treatment, and likely response to specific forms of pharmacological or psychotherapeutic interventions.

THE ASSESSMENT INTERVIEW AND CASE HISTORY

General Considerations

The previously mentioned historical and psychometric considerations indicate that no single correct way exists to conduct an unstructured or semistructured interview. Interviewer style is strongly influenced by theoretical orientation and by practical considerations. Persons strongly influenced by client-centered theories tend to be nondirective and avoid highly structured questions. This is consistent with the underlying belief that persons have the inner ability to change and organize their own behaviors. The goal of a client-centered interview, then, is to create the type of interpersonal relationship most likely to enhance this self-change. In contrast, a behavioral interview is more likely to be based on the assumption that change occurs because of specific external consequences. As a result, behavioral interviews are relatively structured because they are directed toward obtaining specific information that would help to design strategies based on altering external conditions. In addition, different interviewing styles and strategies work well with some clients but may be relatively ineffective with others.

A useful distinction is between a diagnostic interview and one that is more informal and exploratory. The goal of a diagnostic interview is to develop a specific diagnosis, usually based on the multiaxial *DSM-IV* model (see Othmer & Othmer, 1994; R. Rogers, 1995; Sommers-Flanagan & Sommers-Flanagan, 1999). This might follow a five-step process in which the clinician develops diagnostic clues, considers these in relation to diagnostic criteria, takes a psychiatric history, and, based on this information, develops a multiaxial diagnosis with corresponding estimates of prognosis (Othmer & Othmer, 1994). Such an interview is likely to be directive with a careful consideration of inclusion and exclusion criteria for different disorders. It is most likely to occur in a psychiatric or general medical setting. In contrast, many practitioners do not believe in the value of formal diagnosis and, accordingly, do not pursue a formal *DSM-IV* (1994) diagnosis. They might be more concerned with areas such as a client's coping style, social supports, family dynamics, or the nature of their disability. As such, their interviews might be less directive and more flexible. Again, neither style is right or wrong, but instead, may be appropriate and effective in one context (or client), whereas it is ineffective or inappropriate in another context.

Often, interviewers might wish to construct a semistructured interview format by listing in sequence the types of questions they would like to ask the person. To construct such a list, interviewers might consult Table 3.1 to note possibly relevant areas. Each of these areas might then be converted into specific questions. For example, the first few areas might be converted into the following series of questions:

- “What are some important concerns that you have?”
- “Could you describe the most important of these concerns?”

Table 3.1 Checklist for an assessment interview and case history

<i>History of the Problem</i>	
Description of the problem	Intensity and duration
Initial onset	Previous treatment
Changes in frequency	Attempts to solve
Antecedents/consequences	Formal treatment
<i>Family Background</i>	
Socioeconomic level	Cultural background
Parent's occupation(s)	Parent's current health
Emotional/medical history	Family relationships
Married/separated/divorced	Urban/rural upbringing
Family constellation	
<i>Personal History</i>	
Infancy	
Developmental milestones	Early medical history
Family atmosphere	Toilet training
Amount of contact with parents	
Early and Middle Childhood	
Adjustment to school	Peer relationships
Academic achievement	Relationship with parents
Hobbies/activities/interests	Important life changes
Adolescence	
All areas listed for early and middle childhood	Early dating
Presence of acting out (legal, drugs, sexual)	Reaction to puberty
Early and Middle Adulthood	
Career/occupational	Marriage
Interpersonal relationships	Medical/emotional history
Satisfaction with life goals	Relationship with parents
Hobbies/interests/activities	Economic stability
Late Adulthood	
Medical history	Reaction to declining abilities
Ego integrity	Economic stability
<i>Miscellaneous</i>	
Self-concept (like/dislike)	Somatic concerns (headaches, stomach-aches, etc.)
Happiest/saddest memory	Events that create happiness/sadness
Earliest memory	Recurring/noteworthy dreams
Fears	

- “When did the difficulty first begin?”
- “How often does it occur?”
- “Have there been any changes in how often it has occurred?”
- “What happens after the behavior(s) occurs?”

Because clients vary regarding their personal characteristics (age, educational level, degree of cooperation) and type of presenting problem (childhood difficulties, legal problems, psychosis), the questions necessarily need to vary from person to person. Furthermore, any series of questions should not be followed rigidly, but with a certain degree of flexibility, to allow exploring unique but relevant areas that arise during the interview.

Good interviewing is difficult to define, partly because different theoretical perspectives exist regarding clinician-client interaction. Furthermore, clinicians achieve successful interviews not so much by what they do or say, but by making sure they express the proper attitude. Whereas clinicians from alternative theoretical persuasions might differ regarding areas such as their degree of directiveness or the type of information they should obtain, they would all agree that certain aspects of the relationship are essential (Patterson, 1989). These include the interviewer’s expression of sincerity, acceptance, understanding, genuine interest, warmth, and a positive regard for the worth of the person. If clinicians do not demonstrate these qualities, they are unlikely to achieve the goals of the interview, no matter how these are defined.

Patient ratings of the quality of interviews have been found to be dependent on the extent to which interviewers can understand the patient’s emotions and detect emotional messages that are only partially expressed, particularly as these emotions are likely to be indirect and conveyed through nonverbal behaviors. This is especially relevant in clinical interviews that focus on a client’s personal difficulties. Typically, words are inadequate to accurately describe problem emotions, so interviewers must infer them from paraverbal or nonverbal expression. This is highlighted by the assumption that nonverbal aspects of communication are likely to be a more powerful method of conveying information. For example, eye contact is most likely to convey involvement; rigidity of posture might suggest client defensiveness; and hand movements often occur beyond the person’s conscious intent, suggesting nervousness, intensity, or relaxation. Mehrabian (1972) has supported this perspective with his estimates that the message received is 55% dependent on facial expression, 38% by tone, and only 7% by the content of what is said.

Interviewers vary in the extent to which they take notes during the interview. Some argue that note taking during an interview might increase a client’s anxiety, raise questions regarding anonymity, increase the likelihood that he or she will feel like an object under investigation, and create an unnatural atmosphere. In contrast, many interviewers counter these arguments by pointing out that a loss of rapport rarely results solely from note taking during the interview, assuming, of course, that the interviewer can still spend a sufficient amount of time attending to the client. Ongoing note taking is also likely to capture more details and result in less memory distortion than recording material after the interview has been completed. Thus, an intermediate amount of note taking during the interview is recommended. If the interview is audiotaped or videotaped, the reasons for this procedure need to be fully explained, along with the assurance of

confidentiality and the procuring of a signed agreement. Although audiotape or videotape recording is often awkward at first, usually the interviewer and client quickly forget that it is occurring.

Interview Tactics

Numerous tactics or types of statements have been proposed and studied. These include the clarification statement, verbatim playback, probing, confrontation, understanding, active listening, reflection, feedback, summary statement, random probing, self-disclosure, perception checking, use of concrete examples, and therapeutic double binds. Additional relevant topics are the importance of eye contact, self-disclosure, active listening, and touch. These areas are beyond the scope of this chapter, but the interested reader is referred to excellent discussions by Cormier and Cormier (1998), Sommers-Flanagan and Sommers-Flanagan (1999), Sattler (2002), and Zuckerman (2000). The most relevant skills for interviewing do not come so much from memorizing interviewing tactics, but develop from reviewing actual live or taped interview sessions. However, several important tactics of interviewing are described because they provide a general interviewing strategy.

Preliminaries

During the initial phase of the interview, practitioners need to ensure that they deal adequately with the following issues:

1. Organize the physical characteristics of the interview situation so that the room looks used but not untidy; lighting is optimal; seating is arranged so that the interviewer and client are neither too close nor too far and so that eye level is approximately equal.
2. Introduce yourself and indicate how you prefer to be addressed (Doctor, first name, etc.) and clarify how the client prefers to be addressed.
3. State the purpose of the interview, check the client's understanding of the interview, and clarify any discrepancies between these two understandings.
4. Explain how the information derived from the interview will be used.
5. Describe the confidential nature of the information, the limits of confidentiality, and special issues related to confidentiality (e.g., how the information might be obtained and used by the legal justice system). Further, explain that the client has the right not to discuss any information he or she does not wish to disclose. If the information will be sent to other persons, obtain a signed release of information.
6. Explain the role and activities you would like the client to engage in, the instruments that are likely to be used in the assessment, and the total length of time required. In some circumstances, this may be formalized into a written contract (Handelsman & Galvin, 1988).
7. Make sure that any fee arrangements have been clarified, including the hourly rate, total estimated cost, the amount the client versus a third party is likely to need to pay, and the interval between billing and the expected payment.

With the possible exception of fee arrangement (item 7), the preceding issues should be handled by a mental health practitioner rather than a secretary or receptionist. Covering these areas during the preliminary stages of the interview is likely to reduce the likelihood of miscommunications and later difficulties.

Directive versus Nondirective Interviews

The degree to which clinicians choose to be structured and directive during an interview depends on both theoretical and practical considerations. If time is limited, the interviewer needs to be direct and to the point. The interviewer will use a different approach for assessing a person who has been referred and will be returning to the referring person than for a person before conducting therapy with him or her. An ambiguous, unstructured approach probably makes an extremely anxious person even more anxious, while a direct approach may prove more effective. A passive, withdrawn client also is likely to initially require a more direct question-and-answer style. As stated previously, a less structured style often encourages deeper client self-exploration, enables clinicians to observe the client's organizational abilities, and may result in greater rapport, flexibility, and sensitivity to the client's uniqueness.

Frequently, behavioral interviews are characterized as being structured and directed toward obtaining a comprehensive description of actual behaviors and relevant cognitions, attitudes, and beliefs (see Chapter 4). This is often contrasted with the more unstructured psychodynamic approach, which investigates underlying motivations and hidden dynamics, and assesses information that may not be within the person's ordinary awareness. Typically, these approaches are perceived as competing and mutually exclusive. Haas, Hendin, and Singer (1987) point out that this either/or position is not only unnecessary but unproductive, because each style of interviewing provides different types of information that could potentially compensate for the other's weaknesses. Using both approaches might increase interview breadth and validity. This is similar to basing client descriptions on direct behavioral data (public communication), self-description, and private symbolization (Leary, 1957). Each of these levels may be useful for different purposes, and the findings from each level might be quite different from one another.

Sequence of Interview Tactics

Most authors recommend that interviewers begin with open-ended questions and, after observing the client's responses, use more direct questions to fill in gaps in their understanding (Beutler & Groth-Marnat, 2003; Othmer & Othmer, 1994; Sommers-Flanagan & Sommers-Flanagan, 1999). Although this sequence might begin with open-ended questions, it should typically lead to interviewer responses that are intermediate in their level of directiveness, such as facilitating comments, requesting clarification, and possibly confronting the client with inconsistencies.

An important advantage of open-ended questions is that they require clients to comprehend, organize, and express themselves with little outside structure. This is perhaps the only occasion in the assessment process that makes this requirement of clients, because most tests or structured interviews provide guidance in the form of specific, clear stimuli. When clients are asked open-ended questions, they will be most likely to express significant but unusual features about themselves. Verbal fluency, level of assertiveness, tone of voice, energy level, hesitations, and areas of anxiety can be noted.

Hypotheses can be generated from these observations and further open-ended or more direct questions used to test these hypotheses. In contrast to these advantages, open-ended questions can potentially provide an overabundance of detailed, vague, and tangential information.

Interviewer responses that show an intermediate level of directiveness are facilitation, clarification, empathy, and confrontation. Facilitation of comments maintains or encourages the flow of conversation. This might be accomplished verbally (“Tell me more . . .,” “Please continue . . .”) or nonverbally (eye contact, nodding). These requests for clarification might be used when clients indicate, perhaps through subtle cues, that they have not fully expressed something regarding the topic of discussion. Requests for clarification can bring into the open material that was only implied. In particular, greater clarification might be achieved by requesting the client to be highly specific, such as asking him or her to provide concrete examples (a typical day or a day that best illustrates the problem behavior). Empathic statements (“It must have been difficult for you”) can also facilitate client self-disclosure.

Sometimes interviewers might wish to confront, or at least comment on, inconsistencies in a client’s information or behavior. Carkhuff (1969) has categorized the potential types of inconsistencies as being between what a person is versus what he or she wants to be, what he or she is saying versus what he or she is doing, and between the person’s self-perception versus the interviewer’s experience of the person. A confrontation might also challenge the improbable content of what he or she is reporting (“tall” stories).

The purpose of confrontations during assessment is to obtain more in-depth information about the client. In contrast, therapeutic confrontations are used to encourage client self-exploration and behavior change. If a practitioner is using the initial interview and assessment as a prelude to therapy, this distinction is less important. However, a confrontational style can produce considerable anxiety, which should be created only if sufficient opportunity exists to work through the anxiety. Usually, a client is most receptive to confrontations when they are posed hypothetically as possibilities to consider rather than as direct challenges. Confrontations also require a sufficient degree of rapport to be sustained; unless this rapport is present, confrontations probably result in client defensiveness and a deterioration of the relationship.

Finally, direct, close-ended questions can be used to fill in gaps in what the client has stated. Thus, a continual flow can be formed between client-directed or client-organized responses and clinician-directed responses. This sequence, beginning with open-ended questions, then moving to intermediately structured responses (facilitation, clarification, confrontation), and finally ending in directive questions, should not be rigid but should vary throughout the interview.

Comprehensiveness

The basic focus of an assessment interview should be to define the problem behavior (nature of the problem, severity, related affected areas) and its causes (conditions that worsen or alleviate it, origins, antecedents, consequences). Interviewers might wish to use a checklist, such as the one in Table 3.1, to ensure they are covering most relevant areas. In using such a checklist, the interviewer might begin with a general question, such as “How were you referred here?” or “What are some areas that concern you?” Observations and notes can then be made about the way the client organizes his or her

responses, what he or she says, and the way he or she says it. The interviewer could use facilitating, clarifying, and confronting responses to obtain more information. Finally, the interviewer could review the checklist on family background to see if all relevant areas were covered sufficiently. If some areas or aspects of areas weren't covered, the interviewer might ask direct questions, such as "What was your father's occupation?" or "When did your mother and father divorce?" The interviewer could then begin the same sequence for personal history related to infancy, middle childhood, and so on. Table 3.1 is not comprehensive, but is intended as a general guide for most interview situations. If practitioners generally evaluate specific client types (child abuse, suicide, brain-impaired), this checklist may need additional guidelines and/or be used as an adjunct to commercially available structured interviews, such as the Personality Disorder Examination (Loranger, 1988), Neuropsychological Status Examination (Schinka, 1983), or Lawrence Psychological-Forensic Examination (Lawrence, 1984).

Avoidance of "Why" Questions

It is best to avoid "why" questions because they are likely to increase client defensiveness. A "why" question typically sounds accusatory or critical and thus forces the client to account for his or her behavior. In addition, clients are likely to become intellectual in this situation, thereby separating themselves from their emotions. An alternative approach is to preface the question with either "What is your understanding of . . ." or "How did it occur that . . ." rather than "why?" These options are more likely to result in a description rather than a justification and to keep clients more centered on their emotions.

Nonverbal Behaviors

Interviewers should also be aware of their own as well as their clients' nonverbal behaviors. In particular, interviewers might express their interest by maintaining eye contact, being facially responsive, and attending verbally and nonverbally, such as through occasionally leaning forward.

Concluding the Interview

Any interview is bound by time constraints. An interviewer might help to ensure observance of these constraints by alerting the client when only 5 or 10 minutes remain until the arranged completion of the interview. This allows the client or interviewer to obtain final relevant information. There should also be an opportunity for the client to ask any questions or provide comments. At the end of an interview or assessment session, the interviewer should summarize the main themes of the interview and, if appropriate, make any recommendations.

MENTAL STATUS EXAMINATION

The mental status exam was originally modeled after the physical medical exam; just as the physical medical exam is designed to review the major organ systems, the mental status exam reviews the major systems of psychiatric functioning (appearance, cognitive function, insight, etc.). Since its introduction into American psychiatry by Adolf

Meyer in 1902, it has become the mainstay of patient evaluation in most psychiatric settings. Most psychiatrists consider it as essential to their practice as the physical examination is in general medicine (Rodenhauer & Fornal, 1991).

A mental status examination can be used as part of a formal psychological assessment for a variety of reasons. A brief mental status examination might be appropriate before assessment to determine the appropriateness of more formal psychological testing. If, for example, a patient was unable to determine where he or she was and had significant memory impairments, testing with most instruments might be too difficult and could thereby result in needless distress. Such a screening might also be used to determine basic case management issues such as hospitalization or placing the patient under close observation. A mental status examination can also be used as part of an assessment using formal psychological tests. The “raw” data from the exam can be selectively integrated with general background information to present a coherent portrait of the person and assist in diagnosis.

Despite its popularity among psychiatrists, this form of interviewing is not typically used by psychologists, partly because many areas reviewed by the mental status exam are already covered during the assessment interview and through the interpretation of psychological test results. Many psychological tests cover these areas in a more precise, in-depth, objective, and validated manner with scores being compared to appropriate norms. A client’s appearance, affect, and mood are usually noted by attending to behavioral observations. A review of the history and nature of the problem is likely to pick up areas such as delusions, misinterpretations, and perceptual disorders (hallucinations). Likewise, interview data and psychological test results typically assess a client’s fund of knowledge, attention, insight, memory, abstract reasoning, and level of social judgment. However, the mental status examination reviews all of the preceding areas in a relatively brief, systematic manner. Furthermore, there are situations, such as intakes in an acute medical or psychiatric hospital, where insufficient time is available to evaluate the client with psychological tests.

Numerous sources in the psychiatric literature provide thorough guidelines for conducting a mental status exam (Crary & Johnson, 1981; H. Kaplan & Sadock, 2001; Othmer & Othmer, 1994; Robinson, 2001; Sommers-Flanagan & Sommers-Flanagan, 1999), and R. Rogers (1995) has provided a review of the more structured mental status exams. This literature indicates that practitioners vary widely in how they conduct the mental status examination. The most unstructured versions involve merely the clinician’s use of the mental status examination as a set of general guidelines. The more structured versions range from comprehensive instruments that assess both general psychopathology and cognitive impairment to those that focus primarily on cognitive impairment. For example, the comprehensive North Carolina Mental Status Examination (Ruegg, Ekstrom, Evans, & Golden, 1990) includes 36 items that are rated on a three-point scale (not present, slight or occasional, marked or repeated) to cover the important clinical dimensions of physical appearance, behavior, speech, thought processes, thought content, mood, affect, cognitive functioning, orientation, recent memory, immediate recall, and remote memory. Another similar comprehensive instrument is the Missouri Automated Mental Status Examination Checklist (Hedlund, Sletten, Evenson, Altman, & Cho, 1977), which requires the examiner to make ratings on the following nine areas of functioning: general

appearance, motor behavior, speech and thought, mood and affect, other emotional reactions, thought content, sensorium, intellect, and insight and judgment. The checklist includes 119 possible ratings, but the examiner makes ratings in only those areas he or she judges to be relevant.

Despite extensive development, the more comprehensive mental status examinations have not gained wide acceptance. In contrast, the narrower structured mental status examinations that focus more exclusively on cognitive impairment are used quite extensively. One of the most popular has been the Mini Mental Status Examination (Folstein, Folstein, & McHugh, 1975). It comprises 11 items designed to assess orientation, registration, attention, calculation, and language. It has excellent interrater and test-retest reliabilities (usually well above .80), correlates with WAIS IQs (.78 for verbal IQ), and is sensitive to global and left hemisphere deficits (but not right hemisphere impairment; R. Rogers, 1995; Tombaugh, McDowell, Kristjansson, & Hubley, 1996). Clinicians who wish to develop knowledge and skills in conducting mental status examinations are encouraged to consult the preceding sources.

The following descriptions of the typical areas covered serve as a brief introduction to this form of interviewing. The outline is organized around the categories recommended by Crary and Johnson (1981), and a checklist of relevant areas is included in Table 3.2. Interviewers can answer the different areas on the checklist either during or after a mental status examination. The tabled information can then be used to answer relevant questions relating to the referral question, to help in diagnosis, or to add to other test data. Such a checklist is important because clinicians not using similar checklists have been found to frequently omit crucial information (Ruegg et al., 1990).

General Appearance and Behavior

This area assesses material similar to that requested in the “behavioral observations” section of a psychological report (see Chapter 15). A client’s clothing, posture, gestures, speech, personal care/hygiene, and any unusual physical features such as physical handicaps, tics, or grimaces are noted. Attention is given to the degree to which his or her behavior conforms to social expectations, but this is placed in the context of his or her culture and social position. Additional important areas are facial expressions, eye contact, activity level, degree of cooperation, physical attractiveness, and attentiveness. Is the client friendly, hostile, seductive, or indifferent? Do any bizarre behaviors or significant events occur during the interview? In particular, speech might be fast or slow, loud or soft, or include a number of additional unusual features. Table 3.2 includes a systematic checklist of relevant areas of behavior and appearance.

Feeling (Affect and Mood)

A client’s *mood* refers to the dominant emotion expressed during the interview, whereas *affect* refers to the client’s range of emotions. This is inferred from the content of the client’s speech, facial expressions, and body movements. The type of affect can be judged according to variables such as its depth, intensity, duration, and appropriateness. The client might be cold or warm, distant or close, labile, and, as is characteristic of

Table 3.2 Format for mental status and history

Name _____		Observer's Name _____		
		No Data	Present	Absent
APPEARANCE		1. unkempt, unclean, disheveled..... 2. clothing and/or grooming atypical..... 3. unusual physical characteristics.....		
COMMENTS RE APPEARANCE:				
BEHAVIOR	Posture	4. slumped..... 5. rigid, tense.....		
	Facial Expression Suggests	6. anxiety, fear, apprehension..... 7. depression, sadness..... 8. anger, hostility..... 9. absence of feeling, blandness..... 10. atypical, unusualness.....		
	General Body Movements	11. accelerated, increased speed..... 12. decreased, slowed..... 13. atypical, unusual..... 14. restlessness, fidgetiness.....		
	Speech	15. rapid speech..... 16. slowed speech..... 17. loud speech..... 18. soft speech..... 19. mute..... 20. atypical quality, slurring, stammer.....		
BEHAVIOR	Therapist-Patient Relationship	21. domineering, controlling..... 22. submissive, overly compliant, dependent.. 23. provocative, hostile, challenging..... 24. suspicious, guarded, evasive..... 25. uncooperative, non-compliant.....		
COMMENTS RE BEHAVIOR:				
FEELING (AFFECT AND MOOD)		26. inappropriate to thought content..... 27. increased lability of affect..... predominant mood is: 28. blunted, dull, bland..... 29. euphoria, elation..... 30. anger, hostility..... 31. anxiety, fear, apprehension..... 32. depression, sadness.....		
COMMENTS RE FEELING:				
PERCEPTION		33. illusions..... 34. auditory hallucinations..... 35. visual hallucinations..... 36. other types of hallucinations.....		
COMMENTS RE PERCEPTION:				
THINKING	Intellectual Functioning	37. impaired level of consciousness..... 38. impaired attention span, distractible... 39. impaired abstract thinking..... 40. impaired calculation ability..... 41. impaired intelligence.....		
	Orientation	42. disoriented to person..... 43. disoriented to place..... 44. disoriented to time.....		
	Memory	45. impaired recent memory..... 46. impaired remote memory.....		
	Insight	47. denies presence of psychological problems..... 48. blames others or circumstances for problems.....		
	Judgment	49. impaired ability to make routine decisions..... 50. impaired impulse control.....		
THINKING	Thought Content	51. obsessions..... 52. compulsions..... 53. phobias..... 54. depersonalization..... 55. suicidal ideation..... 56. homicidal ideation..... 57. delusions.....		
	Stream of thought	58. associational disturbance.....		
COMMENTS RE THINKING:				
DIAGNOSIS: _____ as manifested by the following M.S.E. items _____ _____				

Reproduced by permission of MTP Press LTD., Lancaster, England, from "Mental Status Examination" by W. G. Crary and C. W. Johnson, 1981. In Johnson, C. W., Snibbe, J. R., and Evans, L. A. (Eds.), *Basic Psychopathology: A Programmed Text*, 2nd ed. Lancaster: MIP Press, pp. 55-56.

schizophrenia, his or her affect might be blunted or flattened. The client's mood might also be euphoric, hostile, anxious, or depressed.

Perception

Different clients perceive themselves and their world in a wide variety of ways. It is especially important to note whether there are any illusions or hallucinations. The presence of auditory hallucinations are most characteristic of schizophrenics, whereas vivid visual hallucinations are more characteristic of persons with organic brain syndromes.

Thinking

Intellectual Functioning

Any assessment of higher intellectual functioning needs to be made in the context of a client's educational level, socioeconomic status, and familiarity and identification with a particular culture. If a low level of intellectual functioning is consistent with a general pattern of poor academic and occupational achievement, a diagnosis of intellectual disability might be supported. However, if a person performs poorly on tests of intellectual functioning and yet has a good history of achievement, organicity might be suspected.

Intellectual functioning typically involves reading and writing comprehension, general fund of knowledge, ability to do arithmetic, and the degree to which the client can interpret the meaning of proverbs. Throughout the assessment, clinicians typically note the degree to which the client's thoughts and expressions are articulate versus incoherent. Sometimes clinicians might combine assessments of intellectual functioning with some short, formal tests such as the Bender, with an aphasia screening test, or even with portions of the WAIS-III or WISC-III.

Orientation

The ability of clients to be oriented can vary in the degree to which they know who they are (person), where they are (place), and when current and past events have occurred or are occurring (time). Clinical observation indicates the most frequent type of disorientation is for time, whereas disorientation for place and person occurs less frequently. When disorientation does occur for place, and especially for person, the condition is relatively severe. Disorientation is most consistent with organic conditions. If a person is oriented in all three spheres, this is frequently abbreviated as "oriented X3."

Related to the orientation of clients is their *sensorium*, which refers to how intact their physiological processes are to receiving and integrating information. Sensorium might refer to hearing, smell, vision, and touch and might range from being clouded to clear. Can the client attend to and concentrate on the outside world or are these processes interrupted? The client might experience unusual smells, hear voices, or have the sense that his or her skin is tingling. Sensorium can also refer to the client's level of consciousness, which may vary from hyperarousal and excitement to drowsiness and confusion. Disorders of a client's sensorium often reflect organic conditions, but may also be consistent with psychosis.

Memory, Attention, and Concentration

Because memory retrieval or acquisition requires attention and concentration, these three functions are frequently considered together. Long-term memory is often assessed by requesting information regarding the client's general fund of information (e.g., important dates, major cities in a country, three major heads of state since 1900). Some clinicians include the Information or Digit Span subtests from the WAIS-III/WISC-III or other formal tests of a similar nature. Recall of a sentence or paragraph might be used to assess short-term memory for longer, more verbally meaningful information. In addition, clients' long-term memory might be evaluated by measuring recall of their major life events, and the accuracy of their recall can be compared with objective records of these events (e.g., year graduated from high school, date of marriage). It is often useful to record any significant distortions of selective recall in relation to life events as well as to note the client's attitudes toward his or her memory.

Short-term memory might be assessed by either requesting that clients recall recent events (most recent meal, how they got to the appointment) or by having them repeat digits forward and backward. Again, the WAIS-III/WISC-III Digit Span subtest, or at least a similar version of it, might be used. Serial sevens (counting forward by adding seven each time) can be used to assess how distractible or focused they are. Persons who are anxious and preoccupied have a difficult time with serial sevens as well as with repeating digits forward and, especially, backward.

Insight and Judgment

Clients vary in their ability to interpret the meaning and impact of their behavior on others. They also vary widely in their ability to provide for themselves, evaluate risks, and make plans. Adequate insight and judgment involves developing and testing hypotheses regarding their own behavior and the behavior of others. Clients also need to be assessed to determine why they believe they were referred for evaluation and, in a wider context, their attitudes toward their difficulties. How do they relate their past history to current difficulties, and how do they explain these difficulties? Where do they place the blame for their difficulties? Based on their insights, how effectively can they solve problems and make decisions?

Thinking

A client's speech can often be considered a reflection of his or her thoughts. The client's speech may be coherent, spontaneous, and comprehensible or may contain unusual features. It may be slow or fast, be characterized by sudden silences, or be loud or unusually soft. Is the client frank or evasive, open or defensive, assertive or passive, irritable, abusive, or sarcastic? Consideration of a person's thoughts is often divided into thought content and thought processes. Thought contents such as delusions might suggest a psychotic condition, but delusions may also be consistent with certain organic disorders, such as dementia or chronic amphetamine use. The presence of compulsions or obsessions should be followed up with an assessment of the client's degree of insight into the appropriateness of these thoughts and behaviors. Thought processes such as the presence of rapid changes in topics might reflect flighty ideas. The client might also

have difficulty producing a sufficient number of ideas, include an excessive number of irrelevant associations, or ramble aimlessly.

INTERPRETING INTERVIEW DATA

Interpreting and integrating interview data into the psychological report inevitably involve clinical judgment. Even with the use of structured interviews, the clinician still must determine which information to include or exclude. Thus, all the potential cautions associated with clinical judgment need to be considered (see Chapter 1). This is particularly important because life decisions and the success of later treatment may be based on conclusions and recommendations described in the report.

Several general principles can be used to interpret interview data. The interview is the primary instrument that clinicians use to develop tentative hypotheses regarding their clients. Thus, interview data can be evaluated by determining whether these hypotheses are supported by information outside the interview. Interview data that is supported by test scores can be given greater emphasis in the final report if it is relevant to the referral question. Even material that is highly supported throughout different phases of the interview process should not be included unless it relates directly to the purpose of the referral.

There is a continuum in handling interview information that varies according to the extent the information will be interpreted. On the one hand, the information might be merely reorganized into a chronological history of the person's life. This would emphasize repeating the information in as objective and accurate a manner as possible. This is typically done in the history section of a psychological report. On the other hand, interview data can be considered raw data to be interpreted. It is thus similar to the data from formal psychological tests. It might, therefore, be used to make inferences related to a client's personality, coping style, or mood and affect.

One method of organizing interview information is to use the information to develop a coherent narrative of the person's life. For example, describing how early family patterns resulted in emotionally sensitive areas ("scar" tissue) can be used to help explain current symptom patterns and difficulties in interpersonal relationships. A different sort of history might trace how interest in a vocation was first begun (early childhood daydreams regarding occupations) and how this progressed and developed as the person matured. Yet, another person might present difficulties related to authority figures. Specific details relating to these difficulties might emerge, such as the client's feeling like a martyr and eventually inappropriately expressing extreme anger toward the authority figure(s). A careful review of the client's history might reveal how he or she becomes involved in these recurring relationships and how he or she typically attempts to resolve them. Persons who are frequently depressed might distance themselves from others by their behavior and then be confused about why relationships seem to be difficult. Often, these themes emerge during a carefully conducted interview, yet aspects of the themes (or the entire themes themselves) are not apparent to the interviewee.

Interview data might also be organized around various domains (see further discussion in Chapter 15). A grid can be used to organize these domains. The various domains might be listed on the left side of the grid with the top of the grid listing the sources of data (of which the interview might be one of a variety of sources of information; see

Figure 15.1 in Chapter 15). Domains might include mood and affect, cognitions, level of resistance, symptom patterns, or coping style. This approach treats interview data in much the same manner as data from psychological tests.

There is no one strategy for sensitizing interviewers to the types and patterns of recurring themes they may encounter during interviews. Inevitably, clinical judgment is a significant factor. The accuracy and types of judgments depend on the theoretical perspective of the interviewer, knowledge regarding the particular difficulty the interviewer is investigating, past experience, types of questions asked, and purpose of the interview.

STRUCTURED INTERVIEWS

Standardized psychological tests and structured interviews were developed to reduce the problems associated with open-ended interviews. They both serve to structure the stimuli presented to the person and reduce the role of clinical judgment. Because structured interviews generate objective ratings on the same areas, they have the advantage of making possible comparisons between one case or population and the next. Typically, these interviews vary in their degree of structure, the relative expertise required to administer them, and the extent to which they serve as screening procedures designed for global measurement or as tools used to obtain specific diagnoses.

Before structured interviews could be developed, clear, specific criteria needed to be created relating to symptom patterns and diagnoses. This ideally helped to reduce the amount of error caused by vague guidelines for exclusion or inclusion in different categories (*criterion variance*). These criteria then needed to be incorporated into the interview format and interview questions. *Information variance* refers to the variability in amount and type of information derived from interviews with patients. In most unstructured interviews, information variance is caused by the wide differences in content and phrasing because of factors such as the theoretical orientation of the interviewer. Structured interviews correct for this by requesting the same or similar questions from each client.

The first popular system of specific criterion-based diagnosis was developed by Feighner et al. (1972) and provided clear, behaviorally-oriented descriptions of 16 psychiatric disorders based on the *DSM-II* (1968). Clinicians using the Feighner criteria were found to have an immediate and marked increase in interrater diagnostic reliability. The descriptions of and relevant research on the Feighner criteria were published in Woodruff, Goodwin, and Guze's (1974) book, *Psychiatric Diagnosis*. Several interviews, such as the Renard Diagnostic Interview (Helzer et al., 1981), incorporated the Feighner criteria. Spitzer, Endicott, and Robins (1978) further altered and elaborated the Feighner criteria to develop the Research Diagnostic Criteria. Simultaneous with the development of the Research Diagnostic Criteria, Endicott and Spitzer (1978) developed the Schedule for Affective Disorders and Schizophrenia (SADS), which was based on the new Research and Diagnostic Criteria. When new versions of the *DSM* were published (1980, 1987, 1994, 2000), revisions of previous interviews typically incorporated the most recent *DSM* criteria along with elements of the Feighner criteria and/or the Research Diagnostic Criteria.

As noted earlier, the reliability of structured interviews has been found to vary depending on the specificity or precision of the rating or diagnosis. Whereas the highest

reliabilities have been found for global assessment (presence/absence of psychopathology), much lower reliabilities have generally been found for the assessment of specific types of behaviors or syndromes. Likewise, high reliabilities have been found for overt behaviors, but reliability has been less satisfactory for more covert aspects of the person, such as obsessions, fears, and worries. Reliability also tends to be lower when clinicians are asked to attempt exact estimates regarding behavioral frequencies and for inferences of multifaceted aspects of the person derived from complex clinical judgments.

Most early studies on validity were based on item content (content validity) or degree of accuracy in distinguishing between broad areas of psychopathology (psychiatric/nonpsychiatric). More recent trends have attempted to assess the accuracy of far more specific areas. However, most validity studies have suffered from an absence of clear, commonly agreed-on criteria. Although structured interviews were attempts to improve on previous, imperfect instruments (unstructured interviews, standardized tests), the structured interviews themselves could not be compared with anything better. For example, the “procedural validity” strategy is based on comparing lay interviewers’ diagnoses with diagnoses derived from trained psychiatrists. Although the psychiatrist’s diagnosis may be better than the layperson’s, diagnoses by trained psychiatrists still cannot be said to be an ultimate, objective, and completely accurate standard. Furthermore, there is confusion about whether actual validity is being measured (which would assume psychiatrists’ diagnoses are the true, accurate ones) or merely a version of interrater reliability. At the core of this issue is the very nature of how diagnosis is defined and the degree to which it is actually helpful in treatment (see Beutler & Malik, 2002; Widiger & Clark, 2000).

Future studies need to involve aspects of what has previously been discussed as construct validity. This means looking more carefully at structured interviews in relationship to etiology, course, prognosis, and treatment utility relating to areas such as the appropriate selection of types of treatments and the likelihood of favorable responses to these treatments. Validity studies also need to look at the interaction between and implications of multiple criterion measures, including behavioral assessment, checklists, rating scales, self-report inventories, biochemical indices, and neuropathological alterations.

Since the mid-1970s, there has been a proliferation of structured interviews for a wide range of areas. Clinicians working in specific areas often select structured interviews directed toward diagnosing the disorders they are most likely to encounter. For example, some situations might benefit from using the Anxiety Disorders Interview Schedule (T. Brown, DiNardo, & Barlow, 1994) to make clear distinctions between anxiety disorders and substance abuse, and between psychosis and major affective disorders. Other contexts might be best served by the Eating Disorder Examination (EDE; Z. Cooper & Fairburn, 1987) or the Structured Interview for *DSM-IV* Dissociative Disorders (SCID-D; Steinberg, 1993). Three categories of structured interviews with representative frequently used instruments are included in Table 3.3 and have been extensively reviewed in R. Roger’s (1995) *Diagnostic and Structured Interviewing: A Handbook for Psychologists*. One consideration in selecting these instruments is that, because most structured interviews are undergoing continuous revisions, the most up-to-date research should be consulted to ensure that practitioners obtain the most recently revised versions. The following pages provide an overview of the most frequently used and most extensively researched structured interviews.

Table 3.3 Frequently used structured interviews by categories

I.	Assessment of Axis I disorders
	Schedule of Affective Disorders and Schizophrenia (SADS) and Schedule of Affective Disorders and Schizophrenia for School-Age Children (K-SADS)
	Diagnostic Interview Schedule (DIS) and Diagnostic Interview for Children (DISC)
	Structured Clinical Interview for <i>DSM-III-R</i> (SCID)
	Diagnostic Interview for Children and Adolescents (DICA)
II.	Assessment of Axis II disorders
	Structured Interview for <i>DSM-III</i> Personality Disorders (SIDP)
	Personality Disorder Examination (PDE)
	Structured Clinical Interview for <i>DSM-III-R</i> Personality Disorders (SCID-II)
III.	Focused structured interviews
	Anxiety Disorders Interview Schedule (ADIS)
	Diagnostic Interview for Borderlines (DIB)
	Psychopathy Checklist (PCL)
	Structured Interview for <i>DSM-IV</i> -Dissociative Disorders (SCID-D)
	Structured Interview of Reported Symptoms (SIRS)
	Psychosocial Pain Inventory (PSPI)
	Comprehensive Drinker Profile (CDP)
	Eating Disorder Examination (EDE)
	Structured Interview of Sleep Disorders (SIS-D)
	Substance Use Disorders Diagnostic Schedule (SUDDS)

Schedule for Affective Disorders and Schizophrenia

The SADS (Endicott & Spitzer, 1978) is a clinician-administered, extensive, semistructured interview that has been the most widely used structured interview for clinical research purposes. Although it was originally designed for differential diagnosis between affective disorders and schizophrenia, it has evolved to include a much wider range of symptoms and allows the interviewer to consider many different diagnostic categories. Although a wide range of disorders is considered within the SADS, its primary strength lies in obtaining fine detail regarding different subtypes of affective disorders and schizophrenia (Summerfeldt & Antony, 2002). The interview rates clients on six gradations of impairment from which diagnoses are reached using the clear, objective categories derived from Spitzer et al.'s (1978) Research Diagnostic Criteria (RDC). The SADS is divided into adult versions for current symptoms, occurrence of lifetime symptoms, and degree of change. There is a further version for the assessment of children's difficulties (K-SADS). Two modifications for the SADS have been the inclusion of anxiety disorders (SADS-LA; Fyer, Endicott, Manuzza, & Klein, 1985, 1995) and eating disorders (EAT-SADS-L; Herzog, Keller, Sacks, Yeh, & Lavori, 1992).

Adult Version

The adult version of the SADS (Endicott & Spitzer, 1978) is designed to be administered in two different parts, the first focusing on the client's present illness and the

second on past episodes. This division roughly corresponds with the three different versions of the SADS. The first is the regular version (SADS), the second is the life-time version (SADS-L, which is actually the second half of the SADS), and the third is the SADS-C, which measures changes in the client. The SADS-L is directed toward diagnosing the possible presence of psychiatric disturbance throughout the person's life. The SADS and SADS-L are the most extensively used. Because the questions in the SADS are directed toward current symptoms and those symptoms experienced one week before the illness, it is most appropriate for administration when the client is having current difficulties. In contrast, the SADS-L is most appropriate when there is no current illness. To make accurate ratings, interviewers are allowed to use a wide range of sources (client's family, medical records) and ask a number of different questions. Final ratings are made on a six-point Likert-type scale. Administration involves more than 200 items and takes from 1.5 to 2 hours and should be conducted only by a psychiatrist, psychologist, or psychiatric social worker. The end product is the following eight summary scales:

1. Mood and ideation.
2. Endogenous features.
3. Depressive-associated features.
4. Suicidal ideation and behavior.
5. Anxiety.
6. Manic syndrome.
7. Delusions-hallucinations.
8. Formal thought disorder.

Interrater reliabilities for the specific diagnostic categories have been found to be quite high, with the exception of the Formal Thought Disorder Scale (Endicott & Spitzer, 1978). The low reliability of this scale may have been because few of the patients in the Endicott and Spitzer sample showed clear patterns of disordered thoughts, which resulted in high variability for the ratings. Test-retest reliabilities were likewise good, ranging from .88 for Manic Disorders to .52 for Chronic and Intermittent Depressive Disorder (Spiker & Ehler, 1984). The exception was a low reliability for schizoaffective, depressed (.24), but this was probably because of the small number of patients included in this category, which resulted in limited variance. Using a different and possibly more appropriate statistical method, reliability increased to .84. Overall, the SADS has demonstrated excellent reliability, particularly for interrater and test-retest reliabilities related to current episodes of psychiatric disturbance.

Validity studies have been encouraging because expected relationships have been found between SADS scores and external measures of depression, anxiety, and psychosis. For example, M. H. Johnson, Margo, and Stern (1986) found that relevant SADS measures could effectively discriminate between patients with depression and paranoid and nonparanoid schizophrenia. In addition, the SADS depression measures effectively rated the relative severity of a patient's depression. For example, Coryell et al. (1994) found clear consistency between different levels of depression. The authors suggest that

incremental validity might be increased by having clients referred for a medical examination to screen out physical difficulties that might be resulting in central nervous system dysfunction. The authors also recommend that interviewers try to increase validity by always including the best available information (family history, structured tests, other rating schedules) before making final ratings. The SADS has been used to predict the clinical features, course, and outcome of various disorders, including major depression (Coryell et al., 1994), schizophrenia (Stompe, Ortwein-Swoboda, Strobl, & Friedman, 2000), and bipolar disorder (Vieta et al., 2000). A number of studies has also successfully used the SADS to detect family patterns of schizophrenia (Stompe et al., 2000) and obsessive compulsive disorders (Bienvenu et al., 2000).

Child Version

The SADS for School-Age Children (Kiddie-SADS-P, K-SADS-P; Ambrosini, 2000; Puig-Antich & Chambers, 1978) is a semistructured interview developed for children between ages 6 and 18. The K-SADS has come out in versions to be used in epidemiological research (K-SADS-E), to assess present and lifetime psychopathology (K-SADS-P/L), and present levels of symptomology (K-SADS-P). Although much of the K-SADS is based on research with major depressive disorders of prepubertal children, it also covers a wide range of disorders such as phobias, conduct disorders, obsessive-compulsive disorders, and separation anxiety.

The interview should be administered by a professional clinician who has been trained in the use of the K-SADS and is familiar with *DSM-III-R/DSM-IV* criteria. All versions are administered to both the parent and the child. Any discrepancies between the two sources of information are clarified before final ratings are made. Total administration time is approximately 1.5 hours per informant (3 hours total). The first phase is a 15- to 20-minute unstructured interview in which rapport is developed as well as an overview of relevant aspects of history, including the frequency and duration of presenting symptoms, their onset, and whether the parents have sought previous treatment. This is followed by structured questions regarding symptoms, which are rated on a Likert scale, with 1 representing "not at all" and 7 indicating that they are "extreme." A skip structure is built into the format so that interviewers can omit irrelevant questions. Interviewers are allowed to use their judgment regarding the wording and the type and number of questions. Finally, ratings are made regarding behavioral observations (appearance, attention, affect). Interviewers are also asked to rate the completeness and reliability of the interview and to make a global assessment of pathology (degree of symptomatology and level of impairment).

Test-retest and interrater reliability for the K-SADS has been good with a general trend for each version to have improved reliabilities. Ambrosini (2000), for example, reported that the K-SADS-P/L had test-retest reliabilities ranging from 1.00 (lifetime occurrence of major depression) to .55 (for lifetime occurrence for attention deficit disorder). However, overall reliabilities have been lower for the K-SADS (and K-SADS-III-R) than for the adult SADS, but this is to be expected given the relative changeableness and less well-developed language skills found with children (Ambrosini, Metz, Prabucki, & Lee, 1989; Chambers et al., 1985). Validity studies indicate that relevant K-SADS measures correlated highly with diagnoses for conduct disorders, schizophrenia,

and depression (Apter, Bleich, Plutchik, Mendelsohn, & Tyrano, 1988). Additional expected correlations have been found between SADS measures and ratings of adolescent mood (E. Costello, Benjamin, Angold, & Silver, 1991) and the Child Behavior Checklist (Achenbach & Edelbrock, 1983; Ambrosini, 2000). Finally, follow-up studies on adolescents diagnosed with disorders (i.e., depression) have found a continued risk for later affective difficulties (i.e., Lewinsohn, Rohde, Klein, & Seeley, 1999).

Collectively, the different versions of the SADS provide a thorough, well-organized interview with unparalleled coverage of the subtypes and gradations of the severity of mood disorders. The SADS has also been well accepted in research and clinical settings. It has strong interrater reliability and provides good ratings of symptom severity, measures associated symptoms, includes guidelines for possible malingering, and has strong evidence of convergent validity (see R. Rogers, 1995; Summerfeldt & Antony, 2002). In contrast, its weaknesses include a relatively narrow band of diagnosis compared with some of the other available instruments such as the SCID or DIS. In addition, the diagnoses are based on Research Diagnostic Criteria (RDC) rather than the more recent *DSM-III-R* or *DSM-IV* criteria. This criticism is somewhat moderated, however, by many of the RDC and *DSM-III/DSM-IV* criteria being nearly the same, especially for childhood disorders. Finally, administration and interpretation of the SADS require extensive training (usually a week) as well as a good working knowledge of differences between the SADS/RDC and *DSM-III-R/DSM-IV* criteria.

Diagnostic Interview Schedule

In contrast to the SADS, which is semistructured and requires administration by trained professionals, the DIS (Robins, Helzer, Croughan, & Ratcliff, 1981) is highly structured and was designed specifically by the National Institute of Mental Health (Division of Biometry and Epidemiology) to be administered by nonprofessional interviewers for epidemiological studies (see Helzer & Robins, 1988). It has been updated for the *DSM-III-R* (DIS-III-R; Robins et al., 1989) and the *DSM-IV* (DIS-IV; Robins, Cottler, Bucholz, & Compton, 1996). The latest version (DIS-IV) includes 19 modules with more than 30 Axis I diagnoses and one Axis II diagnosis (antisocial personality). This modular format allows for tailoring various portions of the DIS-IV to the interests of the researcher or clinician. However, clinical judgment is reduced to a minimum by using verbatim wording, specific guidelines, a clear flow from one question to the next, and simple yes-no answers. Thus, the DIS is far more economical to administer than the SADS. Total administration time is 60 to 90 minutes. Studies have generally indicated that results are comparable between trained clinicians and nonprofessional interviewers (Helzer, Spitznagel, & McEvoy, 1987).

Adult Version

The original version of the DIS was derived from the format of the earlier Renard Diagnostic Interview. However, diagnosis for the DIS-IV is based exclusively on *DSM-IV* criteria. Initially, questions are directed toward obtaining information regarding the client's life, and information is also requested regarding more current symptoms based on the past two weeks, past month, past six months, and past year. Specific probe questions distinguish whether a symptom is clinically significant. A total of 470 potential

clinical ratings are made and organized around 24 major categories. Administration time is approximately 60 to 90 minutes.

Computerized administration and scoring programs are available that can generate *DSM-IV*-based diagnoses. However, computer-based diagnoses on early versions of the DIS were found to generate an average of 5.5 possible diagnoses compared with an average of 2.6 for nonstructured interviews (Wyndowe, 1987). Patient acceptance for the computer administration has been found to be high, although the average administration time is somewhat longer than the clinician-interviewed version.

Studies of the reliability and validity of the DIS have been both variable and controversial. Although much of this research was done on pre-DIS-IV versions, the similarity of format and content between the DIS and DIS-IV suggests that much of this earlier research is pertinent. The comparability of diagnosis by professionals and nonprofessionals using the DIS has generally been supported. This suggests that nonprofessionals can effectively use it to help gather data for large epidemiological studies. For example, Robins et al. (1981) found diagnostic agreement between psychiatrists and nonprofessional interviewers to be .69. The sensitivity (percent interviewees correctly identified) of the DIS varied according to type of diagnosis, but had a mean of 75% with a mean specificity (percent noncases correctly identified) of 94%. More recent studies have similarly concluded that the specificity is stronger than its sensitivity (Eaton, Neufeld, Chen, & Cai, 2000; J. Murphy, Monson, Laird, Sobol, & Leighton, 2000). However, data on sensitivity and specificity were based on using psychiatrists' diagnoses as the true index of diagnostic accuracy. The difficulties in considering psychiatrists' ratings as the truly accurate or "gold standard" criterion for validity have already been noted; therefore, it is probably best to consider the preceding data on sensitivity and specificity as forms of interrater reliability rather than concurrent validity. In contrast to this study, Vandiver and Sheer (1991) found somewhat modest median test-retest reliabilities ranging between .37 and .46.

Although many of the DIS ratings between professional and lay interviewers were equivalent, Helzer et al. (1985) found that, when compared with psychiatrists, nonprofessional interviewers tended to overdiagnose major depression. In contrast to Helzer et al. (1987), Folstein et al. (1985) did not find a sufficiently high rate of agreement between diagnoses by a panel of psychiatrists and diagnoses by the DIS to warrant its use in epidemiological studies. Specifically, it was found that the DIS generated more cases of depression and schizophrenia and fewer cases of alcoholism and antisocial personality (Cooney, Kadden, & Litt, 1990; Folstein et al., 1985). Eaton et al. (2000) has noted that false-negative diagnoses for many cases could be attributed mainly to failure by patients to report symptoms based on life crises or medical conditions. In contrast, the DIS has been found to be comparable with other commonly used psychiatric rating devices such as the Psychiatric Diagnostic Interview (Folstein et al., 1985; R. Weller et al., 1985). However, both diagnostic strategies may contain inaccuracies, and it is difficult to tell in which areas these inaccuracies occurred (R. Weller et al., 1985). The DIS has had the greatest difficulty accurately diagnosing borderline conditions and patients in remission, but this is to be expected because these are the most problematic diagnoses for many other assessment strategies (Robins & Helzer, 1994). In contrast, Swartz et al. (1989) were able to find quite respectable sensitivities (85.7%) and specificities (86.2%) for borderline conditions using a DIS borderline index.

Child Version

The Diagnostic Interview Schedule for Children (DISC; Costello, Edelbrock, Duncan, & Kalas, 1984; Shaffer, Fisher, Lucas, Dulcan, & Schwab-Stone, 2000) is similar to the adult version in that it is highly structured and designed for nonprofessional interviewers. It differs in that it is designed to be given as both a child interview (DISC-C) and parent interview (DISC-P). There have also been versions designed for teachers (Teacher DISC), screening (DISC Predictive Scales), young adults (Young Adult DISC), and administrations that can be given by computer or audiotape (Lucas et al., 2001; Shaffer et al., 2000). Ratings are coded as 0 (not true), 1 (somewhat true), or 2 (very often true). *DSM-IV* diagnoses are generated based on the combined ratings for the child and parent interviews. Some of the more problematic diagnoses (autism, pervasive developmental disorder, pica) are based on an interview with the parent only. The entire interview takes an average of 70 minutes per informant and 90 to 120 minutes per patient, but an explicit skip structure can enable some interviews to be somewhat shorter. The most recent modification of the DISC (DISC-IV; Robins et al., 1996; Shaffer et al., 2000) was designed to be compatible with *DSM-IV* and *ICD-10* criteria. The DISC-IV comprises six modules, each of which comprises the major diagnostic clusters (Anxiety, Mood, Disruptive, Substance Use, Schizophrenia, Miscellaneous).

DISC test-retest reliability (one-year interval) for *DSM-IV* diagnoses in a clinical sample was good to adequate with parent ratings having higher reliabilities (.54 to .79) than child interviews (.25 to .92; Shaffer et al., 2000). However, test-retest reliabilities for a community sample were generally quite poor for child interviews (.27 to .64) but adequate for parent interviews (.45 to .68; Shaffer et al., 2000). Children's reliability increased with age, which is expected considering their increase in intellectual abilities, greater memory, and improved language comprehension and expression. In contrast, reliabilities based on ratings from interviews with the parents decreased with the child's age, probably because the parents have progressively less contact with their child.

Research on the validity of the DISC has found that discriminations between psychiatric and pediatric groups were good for children with severe diagnoses and severe symptoms but not for children with mild-to-moderate difficulties (Shaffer et al., 2000). Discriminations based on interviews with parents were generally more accurate than those based on children (E. Costello, Edelbrock, & Costello, 1985). Accuracy was also higher for externalizing than internalizing disorders (Friman et al., 2000). In addition, comparisons between psychiatric and pediatric referrals indicated that psychiatric referrals had more symptom scores and more psychiatric diagnoses than pediatric referrals (E. Costello et al., 1985). The DISC has also been found to identify risk factors for substance abuse (Greenbaum, Prange, Friedman, & Silver, 1991) and to predict behaviors related to conduct and oppositional disorders (Friman et al., 2000). Ratings between DISC and clinician-based diagnosis were moderate to good (.29 to .74 for parent and .27 to .79 for child; Shaffer et al., 2000) in research settings and followed strict diagnostic guidelines. However, there was very poor agreement between DISC and clinician-based diagnosis when the clinicians performed diagnosis in everyday clinical settings (A. L. Jensen & Weisz, 2002). This may reflect not so much a weakness of the DISC itself, but more that there are considerable differences between how diagnosis is achieved in research as opposed to practice contexts. In summary, the DISC has strengths in that it has

good reliability and validity among clinical samples involving parent interviews, especially when the problems are related to externalizing disorders. However, the DISC is more problematic when ratings are based on child interviews, particularly among community samples and for internalizing disorders.

Diagnostic Interview for Children and Adolescents

The Renard Diagnostic Interview (Helzer et al., 1981) inspired both the DIS and the Diagnostic Interview for Children and Adolescents (DICA; Herjanic & Campbell, 1977; Herjanic & Reich, 1982). It has been through several revisions, which have incorporated the different editions of the *DSM* and elements of the DIS (W. Reich, 2000). Similar to the DIS, the DICA has been designed for administration by lay interviewers. The most recent version was published in 1997 and is available in child, adolescent, and parent versions (W. Reich, 2000). The DICA can be administered to children between ages 6 and 17 years. The format is semistructured and primarily organized around different themes, such as behavior at home, behavior at school, and interpersonal relationships with peers. Additional content areas are substance abuse and the presence of syndromes such as anxiety disorders, mania, and affective disorders. Elaborate instructions are given for skipping irrelevant items, and total administration time is between one to two hours. The administration begins with an interview of both the parent and child, which is designed to establish baseline behaviors and to obtain relevant chronological information. The parent is then questioned about the child to determine the possible appropriateness of common *DSM-IV* diagnostic categories. The final step is to administer a "Parent Questionnaire," which requests additional medical and developmental history and addresses possible diagnoses that have not been covered by previous questioning.

Reliability of the DICA has been quite variable. Test-retest reliability has been quite good, mostly ranging between .76 and .90 (Bartlett, Schleifer, Johnson, & Keller, 1991; Earls, Reich, Jung, & Cloninger, 1988). However, test-retest reliability for child (6 to 12) ADHD was low (.32) and oppositional disorder was low to adequate (.46; W. Reich, 2000). Reliability has been found to be lowest for questions that were complex, related to time, and for children with the highest level of functional impairment. In contrast, questions with the highest reliability were related to frequency and to externalizing symptoms (Perez, Ascaso, Massons, & Chaparro, 1998). Most cross-informant (parent-child) agreement related to specific symptoms has been disappointingly low (.19 to .54; Herjanic & Reich, 1982). The highest level of agreement was for the oldest children and the lowest for younger groups (W. Reich, 2000). Whereas mothers reported more behavioral symptoms, children were more likely to report subjective complaints.

Validity studies on the DICA indicate that it can accurately make the somewhat gross distinction between middle- to older-aged children who were referred to a general psychiatric clinic from those referred to a pediatric clinic (Herjanic & Campbell, 1977). However, there was considerable overlap for children between ages six and eight, thus suggesting that a greater possibility of misdiagnosis exists for children in this age range. The DICA was found to be most effective for assessing relationship problems, less effective for academic difficulties, and least effective for assessing school problems, somatic complaints, and neurotic symptoms (Herjanic & Campbell, 1977). In addition,

adolescents diagnosed with depression on the DICA also had corresponding elevations on the Beck Depression Inventory (Martin, Churchard, Kutcher, & Korenblum, 1991). W. Reich (2000) reported that as the genetic similarity of persons diagnosed with bipolar disorder decreased, their level of psychopathology on the DISC correspondingly decreased. In summary, the psychometric properties of the DICA have been variable with more studies needed to substantiate its validity, particularly concurrent validity (R. Rogers, 1995).

Structured Clinical Interview for the *DSM-IV*

The SCID (First, Spitzer, Gibbon, & Williams, 1996, 1997; Spitzer et al., 1987) is a clinician-administered, comprehensive broad-spectrum instrument that adheres closely to the *DSM-IV* decision trees for psychiatric diagnosis. A certain degree of flexibility is built in so that administration can be tailored to different populations and contexts. Thus, slightly different forms are used for psychiatric patients (SCID-In/Patient), outpatients (SCID-Out/Patients), and nonpatients (SCID-Non/Patients). Criticisms that the early version of the SCID had sacrificed clinical information so that it would be more user-friendly for clinicians resulted in a clear, easy-to-use version for clinical contexts (the SCID-Clinical Version; First et al., 1997) and a longer, more in-depth version for research (SCID-I; First, Spitzer, et al., 1996). Whereas these versions of the SCID are directed toward Axis I diagnoses, a separate version has been developed for the diagnosis of Axis II disorders (SCID-II; Spitzer, Williams, Gibbon, & First, 1990). A further variation, the SCID-D (Steinberg, 1993), has been developed using *DSM-IV* criteria for the assessment of dissociative disorders. The SCID and its variations include several open-ended questions as well as a skip structure, which enables the interviewer to branch into new areas dependent on the client's previous responses. Because clinical judgment is essential throughout the interview, it should be administered only by trained professionals. To increase incremental validity, the authors encourage the inclusion of relevant additional data in making final diagnostic decisions.

The SCID, along with its variations, is the most comprehensive structured interview available. As a result, administration time can be considerable even with the inbuilt screening questions and skip structure. Many individual clinicians and treatment sites deal with this by primarily administering the modules they are most concerned with. For example, a treatment center specializing in substance abuse might administer the module for Psychoactive Substance Use Disorders along with the SCID-II when the comorbidity of personality disorders is suspected. Administration time might also be reduced by administering the computerized mini-SCID (First, Gibbon, Williams, & Spitzer, 1996) that has been designed to screen for possible Axis I disorders. In addition, a computerized SCID-II (AutoSCID-II; First, Gibbon, et al., 1996) that can also potentially reduce clinician time is available. Although it can be administered by telephone, this procedure is discouraged given the poor agreement between telephone and face-to-face diagnoses (Cacciola, Alterman, Rutherford, McKay, & May, 1999).

The SCID and its variations have not been subjected to the level of reliability and validity studies as the SADS or DIS. This might be partially because of the considerable breadth of coverage encompassed by the SCID, making it a daunting task to cover all areas. The reliability studies that have been performed have resulted in overall

moderate, but quite variable, test-retest and interrater reliabilities. For example, interrater agreement using the SCID-II for common diagnostic categories ranges between .40 and .86 with a mean of .59 (First, Spitzer, Gibbon, & Williams, 1995). Riskind, Beck, Berchick, Brown, and Steer (1987) found that several difficult-to-distinguish diagnostic categories had relatively good levels of interrater agreement. These included generalized anxiety disorders (.79, 86% agreement), depressive disorders (.72, 82% agreement; Riskind et al., 1987), panic disorders ($k = .86$), and major depression ($k = .81$; J. Reich & Noyes, 1987). Test-retest reliabilities over a two-week interval for psychiatric patients was fair to good (overall weighted kappas = .61) but poor for nonpatients (overall weighted kappas = .37; J. B. Williams et al., 1992).

For the most part, validity studies of the SCID have assumed that *DSM-IV* diagnoses are the benchmark for making comparisons of diagnostic accuracy. Thus, “procedural validity” has often been assumed since the SCID has closely paralleled the diagnostic criteria derived from the *DSM-IV* (R. Rogers, 1995). A representative validity study found good agreement ($k = .83$) between interviewer ratings and cross ratings of interviewer videotapes by two senior psychiatrists (Maziade et al., 1992). Other studies have found considerable diagnostic overlap within Axis I disorders and between Axis I and Axis II disorders (Alnacs & Torgerson, 1989; Brawman-Mintzer et al., 1993). However, evaluating the meaning of this overlap is difficult because the extent to which it is caused by instrument error versus true comorbidity (i.e., the frequent occurrence of anxiety and depression) is difficult to determine. In contrast to these mostly favorable studies, a number of studies have found generally poor agreement between SCID and clinician-based diagnosis (Shear et al., 2000; Steiner, Tebes, Sledge, & Walker, 1995). In summary, the strength of the SCID is its impressive breadth of coverage, use of modules targeted toward specific areas, and close parallel with the *DSM-IV*. Its weaknesses are its wide variation in reliability and its need for further validity studies, particularly relating it to other diagnostic measures.

RECOMMENDED READING

- Cormier, W. H., & Cormier, L. S. (1998). *Interviewing strategies for helpers: Fundamental skills and cognitive behavioral interventions* (4th ed.). Monterey, CA: Brooks-Cole.
- Othmer, E., & Othmer, S. C. (1994). *The clinical interview using DSM-IV. Vol. 1: Fundamentals*. Washington, DC: American Psychiatric Press.
- Rogers, R. (1995). *Diagnostic and structured interviewing: A handbook for psychologists*. Odessa, FL: Psychological Assessment Resources.
- Sommers-Flanagan, R., & Sommers-Flanagan, J. (1999). *Clinical interviewing* (2nd ed.). New York: Wiley.
- Summerfeldt, L. J., & Antony, M. (2002). Structured and semistructured diagnostic interviews. In M. Antony & D. H. Barlow (Eds.), *Handbook of assessment and treatment planning for psychological disorders*. New York: Guilford Press.