A Review of Instruments for Assessing Family History

Timothy P. Melchert
Marquette University, timothy.melchert@marquette.edu


Timothy Melchert was affiliated with Texas Tech University at the time of publication.
A REVIEW OF INSTRUMENTS FOR ASSESSING FAMILY HISTORY

Timothy P. Melchert
Texas Tech University, Lubbock, TX

ABSTRACT

The influence of family history on one's development has long been a focus of psychological theory, research, and practice. In recent years, however, conceptualizations of family influences on development have evolved considerably, and there has also been increased concern about the reliability of individuals' memory for their childhoods in general. Current knowledge regarding these and other issues are applied to reviewing the instruments that have been developed to assess aspects of family history. The complexity of this type of assessment is emphasized, and a variety of problems with the reliability and validity of the currently available instruments are discussed. Suggestions for future research are also offered.

THE INFLUENCE OF family of origin experiences on development has always been one of the major foci of psychological theory, research, and intervention. The importance of these influences is also reflected in standard psychological evaluations where family history information is normally viewed as an essential component of the assessment. Numerous instruments have been developed to gather information regarding various aspects of one's family history, but the reliability and validity of these instruments have received relatively limited critical examination. Therefore, a comprehensive review of these instruments was conducted so that the strengths and weaknesses of this type of measurement can be identified.

There are several important reasons for critically examining the reliability and validity of family history instruments (FHIs) at this time. First, the research regarding families and development has evolved considerably over recent decades, and the focus and scope of family history assessment has changed substantially as a result. In the first half of this century, family assessment focused heavily on mothers' relationships with their children (e.g., Ackerman, 1938). Starting in the 1960s, however, conceptualizations of family influences on children's development began to change dramatically. Kempe et al. (1962) were the first to bring
widespread professional attention to the problem of child physical abuse, and child sexual abuse began receiving widespread attention about 15 years later (e.g., Finkelhor, 1979). Since that time, a large number of additional influences on family functioning and child development have received substantial empirical attention, including the effects of fathers' behavior (e.g., Parke and Sawin, 1976), child physical neglect (e.g., Wolock and Horowitz, 1984), psychological maltreatment (e.g., Garrison, 1987), parental decision-making style (e.g., Minuchin, 1974), parental substance abuse (e.g., Burk and Sher, 1988), parental psychological maladjustment (e.g., Rutter, 1971), spousal abuse (e.g., Straus et al., 1980), family stressors and coping resources (e.g., Garmezy and Rutter, 1983), the effects of divorce and remarriage (Hetherington, 1989), and subcultural differences between families (Sweet and Bumpass, 1987). Even more detailed aspects of the relationship between parenting and child outcomes are now receiving empirical examination (e.g., differences in parent-child relationships across siblings; Anderson et al., 1994). As a result of this research, comprehensive family history information is needed for many clinical and research purposes. Most of the currently available FHIs, however, were not designed to provide this kind of data.

Another reason for conducting a critical evaluation of FHIs is related to the recent controversy regarding recovered memories of childhood sexual abuse. When individuals describe childhood experiences, they are, of course, describing memories they have of their childhoods, and not a complete and perfectly accurate record of long past events (Alpert et al., 1996). The memories that they describe may be reliable in terms of their internal consistency and temporal stability, but they can also deviate significantly from the original events. Even when individuals report their current feelings or judgments about growing up in their original families, their feelings or judgments are nevertheless based on childhood memories of imperfect reliability. Therefore, aspects of the validity of family history data are fundamentally dependent on the reliability of childhood recall. Since the early history of scientific psychology, autobiographical memory has been known to be less than completely accurate (e.g., Bartlett, 1932; Binet, 1900; Ebbinghaus, 1885/1964; Freud, 1899/1962). Distorted memory did not receive widespread attention, however, until the recent recovered memory controversy (e.g., Schacter, 1995). Precise estimates of the accuracy of childhood recall are not yet available, though most memory experts have concluded that autobiographical memory is reasonably accurate for significant experiences (Baddeley, 1990; Barclay, 1986; Brewer, 1994; Brewin et al., 1993; Heuer and Reisberg, 1992; Neisser, 1994; Ross and Conway, 1986). The literature on FHIs has not included a full consideration of these issues, however. The validity of the data collected with these instruments and priorities for future research in this area consequently need to be reevaluated in light of these considerations.

Another reason for reviewing currently available FHIs is that to date there has been no comprehensive review of these instruments. Several family assessment handbooks have
included brief reviews of selected FHIs (Grotevant and Carlson, 1989; Jacob and Tennenbaum, 1988; L'Abate and Bagarozzi, 1993; Straus and Brown, 1978; Touliatos et al., 1990), but there has been no comprehensive review of the psychometric properties and limitations of FHIs. Critically evaluating the strengths and weaknesses of currently available FHIs is also likely to enhance the quality of future instruments.

In part due to the complexity of family history assessment, interviews have been the preferred approach for gathering family history information in mental health practice and in a great deal of research. Interviewers can often gain an understanding of complex family interactions and memory inconsistencies that instruments cannot capture. Nevertheless, instruments offer advantages over interviews, particularly in terms of efficiency and standardization. Interviews are often time consuming and expensive, and can involve a range of factors that affect the reliability of the information obtained (e.g., unsystematic data collection in unstructured interviews, quality of the interviewing relationship, personal and theoretical biases of the interviewer, leading questions). To avoid broadening the scope of this review to include the complex measurement issues associated with interviews, only questionnaire instruments are included in this review. Readers interested in the use of interviews for gathering family history information are referred to Barone and Switzer (1995), Ceci and Bruck (1995), Millar et al. (1992), Pope and Brown (1996), Rogers (1995), Sagi et al. (1994), and Wyatt and Peters (1986).

INSTRUMENTS

Any pencil-and-paper instrument designed to retrospectively measure memories of at least one characteristic of an adult's family history was included in this review. The use of the terms family history and family of origin in this review are meant to refer to all the family forms in which children are raised, including step, adoptive, and foster families in addition to biological families. The terms mother and father are also used broadly to refer to all the types of parents or parent-figures that a child may have. Instruments that inquire about some family of origin experiences, but are primarily designed to measure aspects of ongoing relationships with original family members (e.g., Personal Authority in the Family System: Bray et al., 1984; Children of Alcoholics Screening Test: Jones, 1983) were not included in this review because there are several issues regarding the retrospective nature of family history assessment that are not involved in current family assessment. For the same reason, instruments designed to measure current family functioning, but which have also been used for assessing family of origin functioning, were not included in this review if no research was located that examined the validity of the instrument for use in family history assessment. Computer searches and family assessment handbooks (Grotevant and Carlson, 1989; Jacob and Tennenbaum, 1988; L'Abate and Bagarozzi, 1993; Straus and Brown, 1978; Touliatos et al., 1990) were used to
locate available FHIs, but several instruments were found by searching the reference lists from various journal articles. Consequently, this review is probably not exhaustive. The 34 instruments that were located are listed and described briefly in Table 1. This review is organized around the traditional approaches to test reliability and validity. After the theoretical basis, structure, and format of currently available FHIs are summarized, the reliability of these instruments is reviewed, and this is followed by a discussion of their content-related, criterion-related, and construct-related validity.

**TABLE 1. CHRONOLOGICAL LISTING AND DESCRIPTION OF FAMILY OF ORIGIN INSTRUMENTS**

*Note. IC = Internal consistency; T-R = Test-retest; MMPI = Minnesota Multiphasic Personality Inventory.*

*aWhen an instrument repeats items and scales to assess mother and father responses separately, the number before the slash indicates the number of separate items or scales, and the number after the slash indicates the number of items or scales after they are repeated for both parents.*

*b"+" indicates that the majority of the coefficients for the majority of the scales in an instrument were greater than .85; "−" indicates that the majority of the coefficients for the majority of the scales were less than .75; "±" is indicated for coefficients that fell between the other two ratings; and “NA” indicates that reliability data were not located.*

*cThe Children’s Reports of Parental Behavior Inventory was also subsequently reduced to 112 items by Margolis and Weintraub (1977) and to 60 items by Schludermann and Schludermann (1988). Across these revisions, however, it retained the three scales proposed by Raskin et al. (1971).*

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Study</th>
<th>Number of Items</th>
<th>Scales</th>
<th>Reliability</th>
<th>IC</th>
<th>T-R</th>
<th>Theoretical Basis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yarrow Parents' Inventory</td>
<td>Yarrow, 1946</td>
<td>70</td>
<td>6 (philosophy of authority, parental restrictions, severity of punishments, parent-child rapport, relative responsibility of mother and father in discipline of child, sibling relationship)</td>
<td>NA</td>
<td>NA</td>
<td>Fel’s Parent- Behavior Rating Scales</td>
<td></td>
</tr>
<tr>
<td>Elias Family Adjustment Test</td>
<td>Elias, 1952</td>
<td>114</td>
<td>9 (attitudes toward mother and father, father/mother attitude quotient, oedipal, struggle for independence, parent-child friction, interparental friction, family inferiority, rejection of child, parental qualities)</td>
<td>+</td>
<td>+</td>
<td>Homey-homelessness</td>
<td></td>
</tr>
<tr>
<td>Instrument Name</td>
<td>Author(s)</td>
<td>N</td>
<td>q</td>
<td>Description</td>
<td>±</td>
<td>±</td>
<td>Theoretical Framework</td>
</tr>
<tr>
<td>-----------------</td>
<td>-----------</td>
<td>---</td>
<td>---</td>
<td>-------------</td>
<td>---</td>
<td>---</td>
<td>-----------------------</td>
</tr>
<tr>
<td>MMPI Harris and Lingoes Pd1 Scale</td>
<td>Harris and Lingoes, 1955</td>
<td>9</td>
<td>1</td>
<td>(family discord)</td>
<td>±</td>
<td>±</td>
<td>Descriptive</td>
</tr>
<tr>
<td>(Unnamed)</td>
<td>Grigg, 1959</td>
<td>15</td>
<td>1</td>
<td>(warmth)</td>
<td>NA</td>
<td>NA</td>
<td>Roe’s Personality Theory of Career Choice</td>
</tr>
<tr>
<td>Parental Role Patterns Questionnaire</td>
<td>Slater, 1962</td>
<td>56112</td>
<td>24</td>
<td>(emotional supportiveness &amp; warmth, inhibitory demands &amp; discipline)</td>
<td>±</td>
<td>NA</td>
<td>Descriptive</td>
</tr>
<tr>
<td>Life Interpersonal History Enquiry</td>
<td>Schutz, 1962</td>
<td>54108</td>
<td>612</td>
<td>(inclusion behavior, control behavior, inclusion feelings, control feelings, affection, disapproval)</td>
<td>−</td>
<td>±</td>
<td>FIRO-B</td>
</tr>
<tr>
<td>(Unnamed)</td>
<td>Switzer et al., 1962</td>
<td>50100</td>
<td>24</td>
<td>(rejecting, overdemanding)</td>
<td>+</td>
<td>NA</td>
<td>Roe’s Personality Theory of Career Choice</td>
</tr>
<tr>
<td>Childhood Experiences Rating Scales</td>
<td>Utton, 1962</td>
<td>30</td>
<td>5</td>
<td>(acceptance, direction of criticism, child-centeredness, rapport, affectionateness)</td>
<td>+</td>
<td>NA</td>
<td>Fel’s Parent- Behavior Rating Scales</td>
</tr>
<tr>
<td>Parent-Child Relations Questionnaire</td>
<td>Roe and Siegelman, 1963</td>
<td>130259</td>
<td>1020</td>
<td>(loving, protecting, demanding rejecting, neglecting, casual, symbolic-love reward/punishment, direct-object reward/punishment)</td>
<td>±</td>
<td>NA</td>
<td>Roe’s Personality Theory of Career Choice</td>
</tr>
<tr>
<td>Family Relations Inventory</td>
<td>Brunkan and Crites, 1964</td>
<td>101202</td>
<td>36</td>
<td>(acceptance, concentration, avoidance)</td>
<td>±</td>
<td>+</td>
<td>Roe’s Personality Theory of Career Choice</td>
</tr>
<tr>
<td>Parent-Child Interaction Rating Scale</td>
<td>Heilbrun, 1964</td>
<td>816</td>
<td>12</td>
<td>(nurturance)</td>
<td>NA</td>
<td>NA</td>
<td>Descriptive</td>
</tr>
<tr>
<td>Perception of Parent Behavior Scale</td>
<td>Apperson and McAdoo, 1965</td>
<td>75150</td>
<td>12</td>
<td>(bothering parents)</td>
<td>NA</td>
<td>+</td>
<td>Descriptive</td>
</tr>
<tr>
<td>Data Source</td>
<td>Measure</td>
<td>Code</td>
<td>Form</td>
<td>Subscales</td>
<td>Notation</td>
<td>Research Framework</td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>---------</td>
<td>------</td>
<td>------</td>
<td>-----------</td>
<td>----------</td>
<td>--------------------</td>
<td></td>
</tr>
<tr>
<td>Family Relations Test</td>
<td>Bene and Anthony, 1965</td>
<td>96</td>
<td></td>
<td>7 (mild and strong positive feelings, mild and strong negative feelings, maternal overprotection, parental indulgence, parental competence)</td>
<td>NA</td>
<td>± Descriptive</td>
<td></td>
</tr>
<tr>
<td>Children's Reports of Parental Behavior Inventory</td>
<td>Schaefer, 1965</td>
<td>260520</td>
<td></td>
<td>2652 (extreme autonomy, lax discipline, moderate autonomy, encouraging sociability, positive evaluation, sharing activities/plans/interests, expression of affection, encourages independence, intellectual stimulation, child-centeredness, possessiveness, protectiveness, intrusiveness, suppression of aggression, strictness, punishment, control through guilt, parental direction, nagging, negative evaluation, irritability, rejection, neglect, ignoring)</td>
<td>±</td>
<td>NA Descriptive</td>
<td></td>
</tr>
<tr>
<td>Children's Reports of Parental Behavior Inventory-192c</td>
<td>Raskin et al., 1971</td>
<td>192384</td>
<td></td>
<td>36 (same as above, but three factors to replace original proposed subscales: acceptance vs. rejection, autonomy vs. control, firm control vs. lax control)</td>
<td>±</td>
<td>+ Descriptive</td>
<td></td>
</tr>
<tr>
<td>Parent-Child Questionnaire</td>
<td>Jacobs et al., 1972</td>
<td>60120</td>
<td></td>
<td>612 (benevolent, domineering, overprotective, ineffective, cold, harsh)</td>
<td>+</td>
<td>NA Leary’s Interpersonal Theory</td>
<td></td>
</tr>
<tr>
<td>Family Violence Scale</td>
<td>Bardis, 1973</td>
<td>25</td>
<td></td>
<td>1 (family violence)</td>
<td>+</td>
<td>+ Eclectic Model of Violence</td>
<td></td>
</tr>
<tr>
<td>Parent Behavior Form</td>
<td>Worell and Worell, 1974</td>
<td>135270</td>
<td></td>
<td>1528 (hostile control, rejection, achievement control, strict control, punitive control, lax control, warmth, active involvement, egalitarianism, cognitive independence, curiosity, cognitive competency, conformity, social desirability, irrationality)</td>
<td>−</td>
<td>NA Descriptive</td>
<td></td>
</tr>
<tr>
<td>Family Environment Scale</td>
<td>Moos et al., 1974</td>
<td>90</td>
<td></td>
<td>10 (cohesion, expressiveness, conflict, independence, achievement orientation, intellectual-cultural orientation, active-recreational orientation, moral-religious emphasis, organization, control)</td>
<td>−</td>
<td>± Social Climate Scales</td>
<td></td>
</tr>
<tr>
<td>Instrument (Title)</td>
<td>References</td>
<td>N/ID/ID</td>
<td>Items</td>
<td>coded?</td>
<td>Description</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>------------</td>
<td>---------</td>
<td>-------</td>
<td>--------</td>
<td>-------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clarke Parent-Child Questionnaire</td>
<td>Paitich and Langevin, 1976</td>
<td>58116</td>
<td>816 (aggression to subject, subject’s aggression to parents, interparental aggression, parental competence, affection, strictness, identification, indulgence)</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Michigan Screening Profile of Parenting</td>
<td>Helfer et al., 1978</td>
<td>17</td>
<td>4 (emotional needs met, relationship with parents, expectations of children, coping)</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parental Bonding Instrument</td>
<td>Parker et al., 1979</td>
<td>2550</td>
<td>24 (care, overprotection)</td>
<td>±</td>
<td>±</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schwarz-Getter Interparental Conflict Scale</td>
<td>Schwarz and Zuroff, 1979</td>
<td>37</td>
<td>1 (interparental conflict)</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schwarz-Getter Interparental Influence Scale</td>
<td>Schwarz and Zuroff, 1979</td>
<td>13</td>
<td>1 (decision-making power)</td>
<td>NA</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inconsistency of Love Scale</td>
<td>Schwarz and Zuroff, 1979</td>
<td>1326</td>
<td>12 (inconsistency of love)</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Own Memories of Child-Rearing Experiences</td>
<td>Perris et al., 1980</td>
<td>81164</td>
<td>1530 (abusive, depriving, punitive, shaming, rejecting, overprotective, overinvolved, tolerant, affectionate, performance oriented, guilt engendering, stimulating, favoring siblings, favoring subject, unspecified)</td>
<td>±</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother-Father-Peer Scale</td>
<td>Epstein, 1983a</td>
<td>3070</td>
<td>37 (independence-encouragement vs. overprotection, acceptance vs. rejection, parent idealization, and peer acceptance vs. rejection)</td>
<td>+</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family of Origin Scale</td>
<td>Hovestadt et al., 1985</td>
<td>40</td>
<td>10 (clarity of expression, responsibility, respect for others, openness to others, acceptance of separation and loss, range of feelings, mood and tone, conflict resolution, empathy, trust)</td>
<td>±</td>
<td>+</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: NA = Not applicable, ± = Ambiguous, + = Positive, Descriptive = Descriptive index
Parent Behavior Report

Schludermann and Schludermann, 1988

2356

24 (nurturance, psychological pressure)

NA

NA

Descriptive

Parental Acceptance Rejection Questionnaire

Rohner, 1989

60

4 (warmth/affection, aggression/hostility, neglect/indifference, undifferentiated rejected)

+ NA

Parental Acceptance-Rejection Theory

MMPI-2 FAM

Butcher et al., 1989

25

1 (family problems)

± ±

Descriptive

Family Ritual Questionnaire

Fiese and Kline, 1993

42

7 (dinnertime, weekends, vacations, annual celebrations, special celebrations, religious holidays, cultural traditions)

± ±

Family Rituals

Psychological Maltreatment Inventory

Engels and Moisan, 1994

25

3 (emotional neglect, hostile rejection, isolation)

+ +

Descriptive

Family Background Questionnaire

Melchert and Sayger, 1998

63179

22 (overall level of family functioning, parental responsiveness, acceptance, physical abuse, physical neglect, sexual abuse, educational involvement, control, decision making style, substance abuse, psychological adjustment, chores, expression of affect, parental coalition, child social support, family stressors)

+ +

Descriptive

THEORETICAL BASIS

Judging on the basis of a stated or inferred theoretical rationale, the majority of the FHIs were developed on a descriptive basis, and relatively few are based on specific theories of families or development (see Table 1). The usefulness of instruments that are tied to particular theoretical orientations is often restricted for practitioners and researchers not subscribing to those orientations. The lack of theory-based instruments, however, makes it more difficult to investigate the validity of particular theoretical approaches (e.g., comparing the ability of different approaches to explain variance in outcome variables).
A small number of FHIs were adapted from child versions which had been developed first (Family Relations Test, Childhood Experiences Rating Scales, and Children's Report of Parental Behavior Inventory), and two of the instruments are part of what have been popular test series (the Life Interpersonal History Enquiry is part of the Fundamental Interpersonal Relations Orientation-Behavior [FIRO-B] test series, and the Family Environment Scale is one of the Social Climate Scales). Three other FHIs also incorporated particular assessment approaches in their development. The Family Relations Test was developed as a play therapy instrument (items are printed on cards that are distributed among cut-out figures representing one's family members), and the Elias Family Adjustment Test and the Parental Role Patterns Questionnaire are based on a projective assessment approach. The two subscales developed from subsets of Minnesota Multiphasic Personality Inventory (MMPI) items have been used very little as independent scales, but were included in this review because of their widespread use.

STRUCTURE AND FORMAT

There is a substantial range in the number of items included in the available FHIs (9-520, median = 70). There has also been a trend toward shorter instruments over the years—in the last 15 years only one instrument has been developed with more than 100 items. Most of the instruments use short statements in a first person format and a Likert-type response scale ranging from 3 to 7 points. Over one-half of the instruments repeat their items (or most of their items) for mothers and fathers. Some of the vocabulary used in the earlier FHIs is now outdated, but the readability of the FHIs as a whole is low.

More than one-half of the FHIs have separate scales for mothers and fathers, but several others inquire about one's family as a whole (Childhood Experiences Rating Scales, Family Environment Scale, Schwarz-Getter Interparental Conflict and Interparental Influence Scales, Family of Origin Scale, Family Ritual Questionnaire, the two MMPI subscales, and the Psychological Maltreatment Inventory). The Parental Acceptance-Rejection Questionnaire inquires only about one's mother. Only one instrument includes respondents' individual siblings in the test format (Family Relations Test). Most of the instruments could easily be adapted for use with various types of family structures, but only two of the instruments use language that explicitly acknowledges different family forms (Mother-Father-Peer Scale and Family Background Questionnaire). Not acknowledging different family forms is likely to have a negative impact on the face validity of FHIs for some individuals from nontraditional families. Users of FHIs and those who develop future instruments should be sensitive to this issue so that stereotypes regarding what constitutes a “normal” family are not unintentionally communicated to respondents. Those administering FHIs to respondents from nontraditional families will also need to clarify which parent(s) or parent-figure(s) the respondent should refer to when completing the instrument. The most common problem appears to involve whether a
respondent should refer to his or her original biological parent or a stepparent who joined the family some time later. In these cases, it may be preferable to ask respondents which of their parents (or parent-figures) had the most influence on them, and ask them to refer to those parents when they complete the instrument.

There are significant differences in the time frame inquired about by the various FHIs. Most of the instruments ask respondents about typical experiences they had in their families of origin and do not discriminate between different periods during childhood and adolescence. Only a few of the instruments specify a time reference for responding to the instrument (the Life Interpersonal History Enquiry uses “around age 6”; the Parent-Child Relations Questionnaire uses “the time before you were 12”; the Parent Behavior Form uses age 16; the Schwarz-Getter Interparental Conflict scale, which was developed with college students, uses “the past five years”; the Parental Acceptance-Rejection Questionnaire uses age 7 to 12). Some of the items in several of the FHIs are confusing regarding the time reference that should be used. The clearest example of this is the Life Interpersonal History Enquiry, which uses age 6 as its time reference, but includes items such as “I wanted my father to be more confident about my ability to think critically,” and “I wanted my mother to be more confident that I would succeed in life.” Clearly children’s relationships with their parents change substantially over time, and many children also experience significant changes in family life as the result of parental divorce and many other factors. These changes are obviously also important for various research and clinical purposes. Therefore, future FHI developers should attend more carefully to this issue. Several psychological theories emphasize the impact of earlier rather than later experiences, and this may provide a rationale for focusing on earlier experiences when respondents report significant changes in their family relationships over their childhoods and test administrators have no reason to focus on one time period over another.

RELIABILITY

Reliability data are missing for several of the FHIs. In the case of several instruments, reliability data were not provided when the instruments were first developed, but were later provided by other researchers. Nevertheless, no reliability data were located for seven of the instruments, and both test-retest and internal consistency data were located for only 41% of the instruments.

There is a huge range in the internal consistency coefficients across the individual scales in the available FHIs (−.14 to .98). Many of the instruments, however, show strong internal consistency. There was often substantial variability in internal consistency across individual scales within instruments and sometimes across different studies of the same instrument. Nevertheless, an average rating of the reliability of each instrument is provided in Table 1.
Test-retest data were located for only one-half of the FHIs. The range of the coefficients that were found was .55 to .97. The test-retest intervals used in these studies ranged from 1 week to 18 months. The large majority of the available coefficients indicate adequate temporal stability for most purposes (see Table 1).

CONTENT-RELATED VALIDITY

The procedures used to assess the content-related validity of psychological instruments are often imprecise, and this has also been the case with regard to FHIs. For example, expert judges were used to evaluate the item content for 10 of the FHIs. The adequacy of these evaluations is difficult to assess, however, because the descriptions of the procedures used were usually insufficient. Future researchers who use these procedures should provide specific information regarding the qualifications of the judges, whether they had been given definitions of the scales, whether they sorted the items into independent scales, and how the representativeness of the item content of the scales was established. Even when these procedures are implemented carefully, however, experts often disagree regarding definitions of particular constructs. Consequently, expert disagreement regarding the adequacy of a group of items for measuring some variable or the adequacy of a collection of scales for measuring some broad construct (such as parental caregiving or family of origin functioning) does not necessarily indicate that an instrument is inadequate. Despite these limitations, the content of FHIs is of particular concern because the domain of variables that is usually considered relevant to family history assessment has become larger and more complex in recent years. Several FHIs were designed to measure specific family variables, and addressing the content-related validity of these instruments is relatively straightforward. Most FHIs were designed to measure parental caregiving or family functioning relatively globally, however, and evaluating the adequacy of the content of these scales is much more complicated.

There are several problems with many FHI items that have been common in the history of psychological measurement, but which also compromise their content-related validity. For example, the two FHIs that were based on a projective approach to assessment include items in the third person, such as “Fathers show dissatisfaction with their families” (Elias Family Adjustment Test) and “Some people are very sociable and gregarious” (Parental Role Patterns Questionnaire). This approach was thought to avoid response biases that would prevent respondents' true thoughts and feelings from being revealed, and this was a popular test format when these two instruments were developed. It fell into disfavor, however, primarily due to problems in knowing what these scales were measuring (Becker and Krug, 1965).

There are also significant problems with the clarity of many FHI items. Items from some of the earlier instruments (e.g., “Some parents are helpful in teaching their child to live comfortably with himself”; Parental Role Patterns Questionnaire), as well as more recent ones (e.g., “My
mother is a person who has trusted me”; Parent Behavior Report), provide numerous examples of potentially confusing items. Likert, the parent of the item format commonly used in FHIs, admonished questionnaire developers to word items as “clear, concise, and straightforward statements,” and “above all … each statement must avoid every kind of ambiguity” (Likert, 1932, p. 45). Unclear items render the content of scales more ambiguous as well as decrease the face validity of instruments.

The focus of many FHI items is also unclear. FHIs are often used to assess memories of family members’ behavior, but many FHI items inquire about feelings or judgments instead of behaviors. For example, the Life Interpersonal History Enquiry emphasizes feelings, and includes items such as “I wanted my mother to play with me more” and “I wanted my father to allow me more freedom.” On the other hand, the Parental Bonding Instrument includes several items that require judgments regarding one’s parents, such as “My [mother/father] did not help me as much as I needed,” and “My [mother/father] did not want me to grow up.” The Family Background Questionnaire is one of the most behaviorally specific of the FHIs, and includes items such as “How often were your mother and father physically violent with you?” and “When your mother or father used alcohol or other drugs, how often did their behavior cause problems (in the home, socially, at work, while driving, etc.)?”

Each of these assessment foci are important, though often for different purposes (e.g., memories of behavior are particularly important in forensic cases; feelings and judgments are important in many psychotherapy cases). They also provide the most comprehensive assessment when they are used together. Items focusing on behavior can be easier to interpret than other types of items, though certainly there are differences in how people interpret many parenting behaviors (e.g., the physical punishment of children). Many people also have great difficulty estimating the frequency of parental behavior they did not observe, such as well-hidden substance use or marital conflict. Nevertheless, it is often more difficult to interpret items inquiring about judgments and feelings than it is regarding the recall of behavior itself. Many individuals lack the insight for judging whether their parents “did not want [them] to grow up,” and items such as “I wanted my father to allow me more freedom” are difficult to interpret without also having behavioral data regarding the context being referred to. It is also difficult to know whether respondents are recalling thoughts and feelings they had when they were children when they respond to items such as these, or if they are reporting current judgments and feelings regarding the behavior they remember their parents engaging in.

As a whole, the currently available FHIs tend to emphasize specific behavior relatively little. It must be noted, however, that restricted behavioral specificity is an inherent limitation of FHIs. Assessing factors such as child maltreatment or parental substance abuse often requires behaviorally specific information. To assess factors such as level of parental responsiveness or behavior control, however, respondents often must average over long time periods and varied
situations in order to provide useful estimates regarding the behavior referred to. It would also be very cumbersome and time consuming to complete a FHI that would specifically accommodate unusual family structures, diverse age, ethnic, and socioeconomic differences, or numerous changes in family circumstances. Consequently individuals with atypical childhoods sometimes cannot complete an FHI, and particularly if they experienced so much change and instability that they cannot reasonably average behaviors over substantial time periods (e.g., some individuals have difficulty identifying parent-figures as a result of a long history of multiple foster placements). Despite the limited behavioral specificity inherent in family history assessment, clinical interview assessments of family history are often more behaviorally focused than most FHIs due to the importance of behaviorally specific information in addition to information regarding feelings and judgments.

As noted above, evaluating the adequacy of FHIs for measuring family of origin functioning comprehensively is complicated because there is no consensus regarding the important variables to include in such an assessment. Family functioning and dysfunction are broad labels (similar to “psychological adjustment” or “maladjustment”) which are useful for referring to heterogeneous classes of processes, patterns, signs, or symptoms, but which have no specific definition (Walsh, 1993). Child development research also has not identified all of the important influences on children’s development and their interactions (e.g., the specific effects of behavioral control, parental alcoholism, or being an incest survivor remain persistent controversies; Kendall-Tackett et al., 1993; Rothbaum and Weisz, 1994; Rotunda et al., 1995).

There is widespread agreement regarding certain issues, however. For example, there is generally consensus regarding the importance of evaluating mother and father behavior separately when assessing parenting or family functioning. Estimates of whole-family functioning are useful for some purposes (particularly for measuring some family systems constructs), but many purposes require more specific information regarding interparental differences. Another widely accepted factor is based on the child socialization research which has consistently found that parental responsiveness and acceptance are among the most important influences on psychosocial outcomes in children (Maccoby and Martin, 1983; Rothbaum and Weisz, 1994). The consensus regarding these two issues is also reflected in FHIs, one-half of which include separate mother and father scales for various forms of parental responsiveness and/or acceptance.

It was noted above, however, that a variety of other variables have also received substantial empirical support in recent years regarding their influence on child development. No theoretical approach has integrated all the diverse factors that have been found to significantly affect family functioning and/or child development. It will also take a substantial amount of
research (and comprehensive, detailed assessment instruments) before the relative importance and interaction of these variables are clarified. For example, the specific effects of different styles of behavior control, parental alcoholism, or being an incest survivor remain unclear. The unique effects of these factors cannot be clarified, however, without more comprehensive and detailed examinations of the interactions between behavioral control, parental substance abuse, incest, other important aspects of parenting, and other major influences on family functioning. Until this type of research is done, firm conclusions regarding the important variables to include in a family history assessment cannot be made.

When a comprehensive measure of family of origin functioning is needed, only the Family Background Questionnaire includes scales covering a wide range of the variables that have been identified as potentially important to children’s development. This instrument is lengthy, however, and is too new to have received significant research attention. When a global measure of parental responsiveness is needed, the FHI which stands out in terms of the research attention it has received is the Parental Bonding Instrument. This instrument has been used in dozens of studies in the United States and other countries, and includes reliable measures of parental care (warmth) and overprotection. The more recent versions of the Children’s Reports of Parental Behavior Inventory have also been fairly widely used, and provide somewhat more reliable and behaviorally specific measures of parental responsiveness and aspects of behavioral control. Other instruments that have received significant research attention include the Family Environment Scale, Family of Origin Scale, Mother-Father-Peer Scale, Own Memories of Childrearing Experiences scale, and the Parental Acceptance Rejection Questionnaire.

CRITERION-RELATED VALIDITY

The two main types of criterion-related validity are predictive and concurrent approaches, and the criteria with which test scores can be compared vary in terms of the directness of their relationship to the test variables. In the past, adult mental health status (especially depression) was often believed to be strongly linked to early parent-child relations, and was considered a relatively direct criterion measure for establishing the validity of family history assessments. More recently, however, the etiology of adult psychopathology has been viewed in a more complex manner. Therefore, studies examining family history and adult outcomes are discussed below in the section on discriminant construct-related validity.

Predictive validity refers to the correspondence between test scores taken at one point and a criterion measure taken at a later point. In the case of FHIs, however, the original experiences that would serve as the criterion occurred in the past. Therefore, this type of validity requires a retrospective approach using behavioral observations or other measures of a respondent's original family made when the respondent was a child. The major problem with this type of study is that original records are often unreliable due to nonrepresentative or inadequate
behavioral observations. The use of original records for documenting specific events that can then be compared with adult recall is very useful in the study of the accuracy of memory for consensually validated discrete events. Family history assessment, however, often focuses on complex interactions and relationships where much of the behavior occurs in private, reactions are often highly dependent on context, and patterns develop and sometimes change dramatically over the course of months and years. When records are correlated with recollections of complex family of origin experiences, it cannot be assumed that the records are sufficiently accurate to be used as the criterion—indeed, they may not even be the better predictor. For example, in the famous study by Yarrow et al. (1970), which has often been cited as demonstrating the unreliability of autobiographical recall, there was generally a great deal of inconsistency between mothers' recall of their children's early development and original nursery school records. Many of the comparisons, however, involved behavior that took place primarily in the home (e.g., typical discipline and affection received by the child), and it is likely that the staff who wrote the original records observed very little of that behavior outside their office, clinic, or school. Given the difficulties involved in doing these kinds of investigations, it is not surprising that no studies of this type were located for any of the FHIs.

Evidence for the concurrent criterion-related validity of FHIs has come from two types of studies. The first of these compared ratings of parent-child relationships made on the basis of unstructured interviews that were then compared with the interviewees' FHI scores. Correlations between ratings and scale scores ranged from .48 to .78 in the study by Parker et al. (1979), and ranged from .23 to .46 in the study by Jacobs et al. (1972). The range in the correlations found in these studies is quite wide, and the implications of these studies are unclear. Factors such as distorted memory and response biases may affect both interview and questionnaire responses, and interview data can be subject to a number of additional limitations (see above). Therefore, the validity evidence provided by this approach is limited.

The second type of FHI concurrent validity study used scores obtained from respondents' siblings or parents as the criterion variable. At least a moderate level of association between siblings would be expected because the same household is shared by all members of a family (i.e., the “shared environment” in the terminology of behavioral genetics). A very high correlation between siblings would not be expected, however, because children from the same family have many unique experiences due to factors such as birth order, different gender role socialization for boys and girls, differences in child temperament that can influence parental caregiving, random events experienced by just one sibling (e.g., a significant illness or injury), or being singled out for favoritism or abuse (i.e., the “nonshared environment”; Dunn and Plomin, 1991; Hoffman, 1991). The proportion of siblings' environments that is shared versus non-shared has not yet been established, though in twin studies a higher correlation between monozygotic (MZ) than dizygotic (DZ) twins would be expected because identical twins are treated more similarly by their parents than fraternal twins (Cohen et al., 1977).
Using data collected with several different FHIs, the intercorrelations of scores from members of the same families has been found to range from statistically nonsignificant to high (range = .14 to .78; Fiese and Kline, 1993; McCrae and Costa, 1988; Melchert and Sayger, 1998; Parker, 1983; Schwarz et al., 1985; Schwarz and Mearns, 1989). In the first of two twin studies, Parker (1986) found that MZ twins did not have more highly intercorrelated scores than DZ twins (.70 for the MZ twin, .71 for the DZ twins). Mackinnon et al. (1991), however, did find more highly intercorrelated scores for MZ twins than for DZ twins (for females, range = .59 to .78 MZ, .51 to .73 DZ; for males, range = .47 to .67 MZ, −.03 to .21 DZ). The majority of the coefficients obtained in these studies are reasonably high, but the lower correlations raise questions about the validity of FHIs. Some of the lower correlations, however, appear to have resulted from the use of less reliable instruments with small, homogeneous samples of college students, which is likely to result in a restriction of range and attenuation of associations.

CONSTRUCT-RELATED VALIDITY

Convergent validity. Tests of convergent construct-related validity have been conducted for several FHIs, and the Family Environment Scale and the Children's Reports of Parental Behavior Inventory have been the most common instruments with which new FHIs were compared. Most of the correlations between the older and newer instruments were high (range = .56 to .90; Hovestadt et al., 1985; Rohner, 1989; Schludermann and Schludermann, 1988; Schwarz and Zuroff, 1979; for exceptions, see Brunkan and Crites, 1964; Fiese and Kline, 1993). The primary weakness of these studies involves the lack of firm evidence establishing the validity of the scales used as the criterion measures.

Discriminant validity. Of the types of validity that have been investigated for the various FHIs, the most common has involved the ability of scores to discriminate between groups expected to differ in family history on theoretical grounds. The most common group to which control subjects were compared has been depressed clients. Freud's (1917/1950) hypothesis that adult depression was associated with the loss of a parent or other loved object in childhood became widely accepted over the next 50 years and provided the rationale for making these discriminations. More recently, however, reviewers have concluded that most of the evidence does not support this hypothesis (Bifulco et al., 1987; Crook and Eliot, 1980; Tennant et al., 1980). Nevertheless, a number of other theories ranging from attachment to social learning to family systems also provide rationales for expecting lower family functioning in various clinical groups. On the other hand, biological theories present competing explanations regarding the etiology of psychiatric disorders, and mood-congruent memory hypotheses provide alternative explanations for associations between depression and negative family of origin memories. Consequently, discriminating between clinical and nonclinical subjects is now viewed as a weaker form of validity evidence for FHIs than it once was.

The large majority of studies examining differences between depressed and control subjects have found greater family of origin dysfunction among the depressed subjects (e.g., Abrahams
and Whitlock, 1969; Bifulco et al., 1987; Fiese and Kline, 1993; Harris et al., 1986; Lewinsohn and Rosenbaum, 1987; McCrae and Costa, 1988; Parker (1979), Parker (1983); Perris et al., 1987; Raskin et al., 1971; Schwarz and Zuroff, 1979). Similar patterns of score differences have also been found with other groups, though the strength of the rationales for selecting risk factors for family of origin dysfunction has varied widely. Some of the weakest rationales involved expectations that male and low socioeconomic status (SES) subjects would obtain higher original family violence scores than female and high SES subjects (Bardis, 1973), while stronger rationales were used in studies of adult children of alcoholics and incest survivors (Hovestadt et al., 1985; Melchert, in press). A variety of other groups have been used, including individuals with anxiety disorders (e.g., Arrindell and Van Der Ende, 1984; Hafner, 1988; Leon and Leon, 1990; Parker, 1981), substance abuse problems (e.g., Apperson and McAdoo, 1965; Hovestadt et al., 1985), schizophrenia (e.g., Parker et al., 1982; Baker et al., 1987; Warner and Atkinson, 1988), eating disorders (e.g., Calam et al., 1990; Reeves and Johnson, 1992), various sex-role orientations (Kelly and Worell, 1976), and various parenting practices (e.g., Biringen, 1990; Jacobvitz et al., 1992). Clearly the relationship between family history and adult psychosocial outcomes is complex. Taken as a whole, however, these studies provide generally supportive evidence regarding the validity of FHIs.

The effects of response styles on FHI data has received only limited research attention. Only one FHI includes a validity scale (the Mother-Father-Peer Scale includes Mother and Father Idealization subscales), and social desirability is the only other response style that has been examined in the FHI literature. The content of FHIs generally concerns sensitive and personal matters, and it would be expected that responses may be biased in a socially desirable direction. Correlations between various FHIs and social desirability scales have mostly been nonsignificant and the strongest ones have been weakly positive (range = −.31 to .37; Melchert and Sayger, 1998; Parker, 1979; Schludermann and Schludermann, 1988; Siegelman, 1965). The level of these correlations is encouraging, but a more complete examination of the effect of response styles on FHI scores is warranted because of their potential clinical significance (e.g., defensively minimizing family problems, exaggerating family dysfunction out of malingering or a defensive justification of current problems, logically inconsistent responding reflecting inconsistent family memories). The direction of causality between response styles and family of origin history also needs to be investigated. For example, it is usually assumed that positive correlations between social desirability and psychological measures indicate that social desirability is causing an inflation of scores, but it is also possible that family of origin experiences causally influence the development of this response style.

**Factor analysis.** Factor analyses have been conducted on the majority of FHIs, and factor analyses formed the rationale for the scales in several instruments. The modal number of factors interpreted for the various FHIs was three (range = 1–20). All of these analyses found at least one factor related to parental responsiveness, though a wide variety of other factors were found, depending on the instrument.
Factor analyses of individual instruments have been quite useful for clarifying the structure of those FHIs. For example, the numerous overlapping original scales on the Own Memories of Childrearing Experiences inventory and the Children’s Report of Parental Behavior Inventory were replaced with three factor analytically derived scales (Arrindell et al., 1983; Raskin et al., 1971). Factor analytic findings across FHIs, however, have been less useful for elucidating the important dimensions underlying family of origin functioning. The factors that emerge from an analysis of an instrument depend on the range of items included, and instruments that gather data regarding a restricted domain of constructs will result in a factor structure reflecting only that restricted domain. As this review has found, most of the available FHIs cover a restricted range of family variables. Melchert and Sayger (1998), on the other hand, found that a 20-factor solution corresponded closely to the 22 scales in the Family Background Questionnaire. Factor structures across clinical, gender, or ethnic groups have not been compared for any of the FHIs, however.

**Longer-term stability of childhood memory.** In general, family of origin memories would be expected to be stable over the short-term and the long-term. Of course, childhood memories could be both temporally stable and inaccurate, but they would have to be temporally stable in order for them to be veridical. A phenomenological perspective would emphasize one’s perception and subjective understanding of one’s family experiences rather than the objective accuracy of one’s family of origin memories. Even from this perspective, however, it would be expected that family of origin memories would be temporally stable in the absence of intervening events that would explain a change in the memories. There may be several factors that could account for such changes over long time spans (e.g., processing family of origin issues in psychotherapy, discussing family experiences with family members and relatives, reviewing old photographs or diaries, maturation). Over shorter periods, however, these memories would generally be expected to be fairly stable.

Several studies have examined the longer-term stability of FHI scores in clinical subjects (Abrahams and Whitlock, 1969; Gerlsma et al., 1993; Gerlsma et al., 1994; Gotlib et al., 1988; Parker, 1981), and three studies used nonclinical samples (Gerlsma et al., 1994; Richman and Flaherty, 1987; Wilhelm and Parker, 1990; see also Finlay-Jones et al., 1981). All of the studies using clinical subjects found no significant differences between test and retest scores over periods ranging from 6 to 30 months (recall also did not become more positive as depression remitted in these studies, which does not support the mood-congruent recall hypothesis). The studies using nonclinical subjects and larger sample sizes did find significant differences between test and retest scores. These differences tended to be quite small over a 6- and 7-month period, however (from .79 to .81 in the Richman and Flaherty study; from .78 to .89 in the Gerlsma et al. study). The longest-term stability of FHI scores was examined by Wilhelm and Parker (1990) who retested college students after 5 and 11 years. Test-retest correlations of .67 to .82 were found across the scales used over the 5-year retest period, and correlations of .56 to .72 were found over the 11-year retest period.
NORMATIVE DATA

Data have been provided for eight of the available FHIs that were called “‘normative,” but the representativeness of most of the subject samples used in these studies was quite limited. Most of the data was obtained from college student samples, and in most instances the samples were recruited from only one university. The most adequate normative data is available for the Life Interpersonal History Enquiry (\( N = 5847 \), with some socioeconomic diversity represented). In the case of FHIs, normative data would be especially useful if they were presented for various clinical, nonclinical, and risk groups. For example, it would be useful to have separate group norms for families with histories of parental chemical dependency and physical, sexual, and/or emotional abuse along with norms for families without any of these characteristics. Normative data for various ethnic, socioeconomic, and family structure groups would also be useful. None of these types of data have been presented for any of the FHIs.

DISCUSSION

Family of origin influences on development have been one of the primary foci of psychological theory and practice over the last century, but the reliability and validity of FHIs have received relatively little critical examination. This is the first comprehensive review of FHIs, and it found that many of the available instruments suffer major weaknesses. In fact, for some purposes the weaknesses predominate, and several questions regarding the validity of FHI data have received very little empirical examination. A basic but serious problem with many FHIs is that reliability data are unavailable or they suggest that the instruments are inadequate for applied purposes or even for research purposes. The large variability in coefficients across FHIs, however, suggests that it is the instruments with low reliability that may be inadequate, and not that family of origin memories themselves are highly inconsistent or temporally unstable. Another basic problem with most FHIs for some purposes is their limited coverage of family variables, and only one of the currently available FHIs attempts a comprehensive coverage of family history. There are also significant problems with the clarity and focus of many of the items in several of the instruments.

Many of the other questions raised in this review concern the veridicality of the data obtained with FHIs. These questions, however, have generally received very limited direct examination. Inaccuracies in autobiographical memory have long been of concern in psychology, but they did not receive a great deal of attention until relatively recently. Most memory experts have concluded that autobiographical memory is reasonably accurate (see above), but there is also widespread consensus that human memory involves complex reconstructive processes that often result in distorted memory (Bartlett, 1932; Bonanno, 1990; Freud, 1899/1962; Lindsay...
and Read, 1994; Loftus and Loftus, 1980; Schacter, 1995; Spence, 1982). General conclusions such as these are useful, but they beg a number of specific questions regarding the precise level of accuracy in childhood memory, individual differences in memory reliability, and differences in memory reliability across domains of experience (e.g., memory for discrete events compared with memory for the nature of parent-child relationships).

The majority of the research reviewed above supports the general conclusions arrived at by most memory experts regarding the reliability of autobiographical memory. For example, FHI scores have been found to be reasonably stable over periods of up to several years. Even in samples of clinical subjects whose problems with mood remitted over the course of the study, no significant difference in FHI scores has been found over periods of up to 30 months. The majority of the studies of siblings' FHI scores have also found moderate to high intercorrelations, the range that would be expected from a consideration of their shared and nonshared family environments. Further, individuals from families with risk factors for dysfunction obtained scores indicating significantly lower family functioning than individuals without these risk factors. Unfortunately, more precise conclusions regarding the validity of FHI data are not possible until future research clarifies a number of the questions that were raised in this review.

FUTURE RESEARCH

The most direct examination of the objective accuracy of FHI data involves longitudinal studies that compare reported family history with reliable records of the experiences referred to. As noted above, however, records that are sometimes available often include inadequate and unrepresentative behavioral observation, and consequently cannot provide a reliable criterion with which to compare childhood recall. Similar problems have been encountered in personality research. Mischel (1968), for example, noted that an individual's observed behavior in one situation tends to correlate with observed behavior in another situation or with scores on personality instruments below .30. However, when more reliable behavioral observations of individuals are made over a number of instances and situations, the stability of personality measurement increases greatly (Epstein, 1980). In fact, research has repeatedly found that knowledgeable raters' impressions are equal to or exceed the predictive power of directly recorded behavioral observations (Eaton and Enns, 1986; Moskowitz and Schwarz, 1982; Weinrott et al., 1981). Epstein (1983b) suggested that informants who are knowledgeable about individual subjects engage in a kind of intuitive averaging after they have observed the subjects on many occasions, and their ratings may be more accurate than behavioral observations because they are usually based on a greater amount of observation than behavioral ratings typically are. These findings raise the possibility that individuals' reports of their family histories may be more reliable than recorded behavioral observations. Research
investigating these issues with regard to FHI data is badly needed, though clearly it is difficult to conduct.

There are other important questions regarding the validity of FHI data that do not rely on behavioral observations and are more easily investigated. For example, the internal consistency and temporal stability of family of origin memories could be examined in detail in cases where family of origin issues are a focus of psychotherapy. Clearly clients may change their perception and understanding of family relationships after developing a more complete understanding of their family history--this is a prominent aspect of many approaches to psychotherapy. While clients' judgments and feelings about their family members often appear to change substantially over the course of psychotherapy, perhaps clients' memories of their family members' behavior actually change relatively little. Perhaps the feelings and judgments that they can recall having as children also change very little over the course of therapy.

Multitrait, multimethod designs may also provide important evidence regarding the validity of FHI data. For example, individuals' reports of their own history could be compared with their siblings' reports of their own histories. These in turn could be compared with the original subjects' observations of their siblings' histories, and the siblings' observations of the original subjects' histories. In addition to comparing self-versus other reports, a number of additional hypotheses could be tested with this data. For example, it would be hypothesized that gender-role socialization effects would result in higher correlations between same-sex siblings reporting on their own history than between opposite-sex siblings; that birth order effects would result in higher scores for first-born children on parental responsiveness than for their later-born siblings; that reports of shared family experiences affecting the whole family (e.g., parental substance abuse or psychological maladjustment) would be more highly intercorrelated between siblings than reports of experiences that are more dependent on individual parent-child relationships (e.g., parental responsiveness, control, maltreatment). Research such as this could provide relatively compelling evidence regarding the validity of FHI data.

Another methodology that could provide important validity evidence involves comparisons of FHI scores between groups with risk factors for original family dysfunction. Given the large amount of variability in families with any one of these characteristics, clearer signs of family dysfunction will result from combining these risk factors. For example, it would be hypothesized that greater numbers of risk factors for family of origin dysfunction (e.g., parental substance abuse; child physical, emotional, or sexual abuse; spousal abuse) will result in FHI scores indicating the lowest levels of family functioning compared to groups with fewer risk factors, and groups with no risk factors will score the highest on family of origin functioning.
Independent verification of the presence of these risk factors would significantly strengthen these kinds of studies, but this type of verification is obviously difficult to obtain.

The phenomenology of family of origin memory also needs to be better understood. Though obviously less important than the veridicality of these memories for many purposes, a phenomenological perspective is important in part because perceptions and memories of one's family history that are subjectively believed to be accurate but that are objectively inaccurate can still have a significant influence on psychological functioning (cf. psychodynamic theories, cognitive approaches to psychotherapy, research in social cognition and the self-concept). A phenomenological perspective also highlights the complexity involved in validating family of origin memories. For example, when the recall of parents or other family members is available for comparison purposes, agreement between family members regarding what happened does not necessarily indicate that the recall is veridical because all of the family members may hold inaccurate perceptions of what happened (e.g., “family myths” can be mutually held by all members of a family; Ferreira, 1963; Stierlin, 1973). Inconsistencies between family members' recall also does not necessarily indicate that one person's recall is inaccurate because the family members may hold equivalently inaccurate perceptions of what actually happened (e.g., the individuals involved in some family conflicts appear to be equivalently right and/or wrong; Silverstein, 1997). In fact, in phenomenologically oriented research and therapy, a person's self-report is often assumed to reflect his or her internal subjective world which can only be completely known to that individual, and which is assumed to be accurate for that person. Consequently, there may not exist a meaningful external, objective reality with which these reports can be compared (e.g., Rogers, 1961; see also Spence, 1982, for a discussion of the “narrative” versus “historical” truth of autobiographical memory).

CONCLUSION

This review found that some family history assessment needs can be adequately addressed by some of the currently available FHIs. It was also found, however, that many questions regarding the validity of FHI data have received limited or no empirical examination. In particular, specific conclusions regarding the objective accuracy of the memories assessed by FHIs must await the findings of future research. At the present time, it is clear that some family of origin memories are not veridical, but they are likely to be reasonably accurate in general. Current research does not support discounting family history data in general because they are not completely veridical or accepting them as prima facia valid out of a strong phenomenological perspective. Instead, the complexity of the assessment issues involved needs to be acknowledged, and interpretations of the data must be made cautiously. The reliability and validity of family history assessment has received relatively limited research attention in psychology compared with other important variables such as personality or intelligence. The importance of family history
information to psychological research and practice, however, suggests that the issues raised in this review deserve much more thorough examination as well.

REFERENCES


Ackerman (1938) N.W. Ackerman The unity of the family Archives of Pediatrics, 55 (1938), pp. 51-62


Biringen (1990) Z. Biringen Direct observation of maternal sensitivity and dyadic interactions in the home: Relations to maternal thinking Developmental Psychology, 26 (1990), pp. 278-284
Bonanno (1990) G.A. Bonanno Remembering and psychotherapy Psychotherapy, 27 (1990), pp. 175-186
Epstein (1983b) S. Epstein Aggregation and beyond: Some basic issues in the prediction of behavior Journal of Personality, 51 (1983), pp. 360-391
Ferreira (1963) A.J. Ferreira *Family myths and homeostasis* Archives of General Psychiatry, 9 (1963), pp. 457-463


Harris et al. (1986) T. Harris, G.W. Brown, A. Bifulco *Loss of parent in childhood and adult psychiatric disorder: The role of lack of adequate parental care* Psychological Medicine, 16 (1986), pp. 641-659

Heilbrun (1964) A.B. Heilbrun *Parental model attributes, nurturant reinforcement, and consistency of behavior among adolescents* Child Development, 35 (1964), pp. 151-167


Likert (1932) R. Likert A technique for the measurement of attitude Archives of Psychology, 140 (22) (1932), pp. 5-55

Lindsay and Read (1994) D.S. Lindsay, J.D. Read Psychotherapy and memories of childhood sexual abuse: A cognitive perspective Applied Cognitive Psychology, 8 (1994), pp. 281-338


Parker (1979) G. Parker Parental characteristics in relations to depressive disorders British Journal of Psychiatry, 134 (1979), pp. 138-147


Schaefer (1965) E.S. Schaefer Children's reports of parental behavior: An inventory Child Development, 36 (1965), pp. 413-424


Slater (1962) P.E. Slater Parental behavior and the personality of the child Journal of Genetic Psychology, 101 (1962), pp. 53-68


