Don't Leave HOME Without It: A Review of the Evidence on the Reliability and Validity of the Long and Short Forms of the HOME Inventory, and Discussion of Implications for Further Survey Work

Carrie L. Mariner and Martha J. Zaslow

Child Trends, Inc.

Methods Working Paper # 98.1

This paper begins with a discussion of the origins of the HOME Inventory, noting especially its conceptualization as a summary of environmental risk. The process by which the HOME Inventory was modified and abbreviated for survey use (the HOME-Short Form) is briefly summarized. The paper then provides a summary of the evidence on the reliability (internal consistency, test-retest, and interrater) as well as validity (concurrent and predictive) of both versions of the HOME. Internal consistency of the subscales of the long form are generally quite strong, with a few noteworthy exceptions. The two subscales of the HOME Short-Form show somewhat weaker internal consistency than the subscales of the long form. The Emotional Support subscale consistently shows lower internal consistency than the Cognitive Stimulation subscale. Internal consistency of the Emotional Support subscale of the HOME-Short Form proved particularly problematic when the measure was used in a low income, predominantly minority sample.

While interrater reliability has been documented for the interviewer ratings completed for the HOME long form, there is a noteworthy absence of data on interrater reliability for the short form. This is a serious gap given the differences in training for those generally administering the long and short forms of the HOME. In terms of test-retest reliability, correlations are higher across shorter time periods, but there is evidence of consistency even over longer periods. Test-retest reliability is better for older children. Findings on validity indicate both the long and short form subscales to be strongly related to measures of socioeconomic status (though ceiling effects at upper socioeconomic levels have been documented), and to predict to measures of child intelligence and achievement. There is some evidence of better predictive validity for children from nonminority than from minority families.

Based on this evidence, the paper concludes by noting ways in which a survey measure of parenting and the home environment could be strengthened: (1) the possibility of a survey measure scored as a full continuum measure rather than as a sum of environmental risk factors; (2) the need to clarify the underlying constructs of subscales for a survey measure; (3) the need to extend the content covered by the discipline component of a survey measure of parenting beyond the issue of physical discipline; (4) the need to address the materialistic basis of many of the items pertaining to cognitive stimulation; (5) the need to document interrater reliability in a survey version of the measure; and (6) the need for further careful examination of internal consistency and predictive validity of the short form when used in differing racial/ethnic groups.

Don't Leave HOME Without It: A Review of the Evidence on the Reliability and Validity of the Long and Short Forms of the HOME Inventory, and Discussion of Implications for Further Survey Work[1]

Carrie L. Mariner and Martha J. Zaslow

Methods Working Paper # 98.1

Introduction

Social scientists from many fields believe that the home environments of young children have important effects on children's development. In an attempt to quantify the crucial aspects of the home environment which would identify those children at high risk for later developmental problems, Bradley and Caldwell (1984) designed and later refined the Long Forms of the Home Observation Measure of the Environment, also known as the HOME. The HOME Long Form has several versions, which vary by age of the child, and a HOME Short Form has been developed for use in large-scale surveys. The Short Form contains items that can be quickly completed by the parent and interviewer to accommodate the time constraints of survey research. The various forms of the HOME have become widely used measures of the home environment. In this working paper we describe the HOME instrument and its development, summarize current knowledge about its reliability and validity, and discuss the

[1]This paper was completed as part of the work on NICHD Grant No. R01 HD31056. Readers should note that the literature review on the reliability and valdity data was completed in 1995. Further work may have been completed subsequently. The authors thank Kristin A. Moore, Kathryn Tout, and Tamara Halle for helpful feedback on the manuscript, and Julie A. Floryan for timely assistance in finalizing the document.

implication of these findings for strengthening measures of parenting and the home environment which are appropriate and reasonable for use in surveys.

Instrument Development

The Origins of the HOME Long Form

When designing each of the age-appropriate versions of the HOME, the designers began with a review of the child development research literature and theory (Bradley, Caldwell & Rock, 1988; Bradley et al., 1992). They also consulted professionals who worked with children to discuss the appropriateness of the items and the topics which should be covered (Bradley et al., 1992). Starting with a large number of items (in the neighborhood of 200 for each version), the authors used data from successive field tests to remove items which did not work well in factor analysis or item analysis (Bradley, Caldwell & Rock, 1988; Bradley et al., 1992). Subscales were formed based mainly on factor analysis, with some subscales formed by item content despite lack of a factor (for a more complete description of the subscales, see Bradley et al., 1994). Subscales were intended to represent correlated yet distinct aspects of a child's environment. Originally, the authors intended only to use observational data, based on a visit to the home (Bradley & Caldwell, 1979). It became clear, however, that relying solely on observation narrowed the focus of the measure to behaviors and materials which were available to be observed in the one hour visit. Important topics such as daily routines could not be addressed. Therefore, the authors added some items which would need to be answered through informal, conversational interviewing of the mother, with input from the child allowed on the Adolescent version (Bradley, personal communication, 1995). In its current form, the HOME Long Form does not

distinguish between observational and interview questions, but leaves administrators of the instrument to decide on a case by case basis whether they need to ask the mother about each item.

In designing and using the HOME scales, Bradley and Caldwell have mentioned several motivations for the scales' development. According to one article, "[t]he original intent... was to more accurately assess the quality and quantity of experiences available in a child's home than was possible with measures such as social class or socioeconomic status designations" (Bradley & Caldwell, 1984a, p.315). However, in an article more contemporaneous to the scales' development, the authors describe the HOME as "...primarily intended as a screening instrument" (Bradley & Caldwell, 1979, p. 237). Indeed, although the HOME is most accurately described as a measure of risk, it has been used by social scientists from various fields in numerous ways: to investigate environmental correlates of infant characteristics (Affleck et al., 1982; Berlin et al., 1995; Parks & Bradley, 1991), to predict cognitive development (Bradley & Caldwell, 1980; Bradley et al., 1989; Moore & Snyder, 1991; Ramey, Yeates & Short, 1984; Van Doorninck et al., 1981; Widmayer et al., 1990; Yeates et al., 1983), to measure the competencies of different groups of parents (Coll, Hoffman & Oh, 1987), as an index of protective factors (Bradley et al., 1994), and as a general measure of parenting (Klebanov, Brooks-Gunn & Duncan, 1994; Metzl, 1980; Sugland et al., 1995), to name just a few.

Each item on the HOME is scored in a binary choice format (Bradley & Caldwell, 1984). Because the scale was designed as a measure of risk, cutoffs for many of the items are set so that a score of zero on an item indicates the lack of one adequately stimulating or supportive aspect of the environment. For example, if the interviewer observes any warmth in the mother's voice

when she is talking to the child, this item is scored one. A lack of warmth in the mother's voice when talking to her own child in her own home is likely to indicate a fairly serious lack of warmth in the child's environment. These adequate versus inadequate cutoffs are reflected in the high rate of positive answers to many of the items (Bradley & Caldwell, 1979; Bradley et al., 1992). It should be noted that this format does not differentiate between adequate and more than adequate environments, such as a little warmth versus a lot of warmth.

The Origin of the Short Form

During the mid 1980s, the HOME was abbreviated and adapted for use in the National Longitudinal Survey of Youth-Child Supplement (NLSY-CS), as part of a broader attempt to include in this survey measures related to children's development. The HOME-Short Form was first used in the 1986 wave of the NLSY-CS, and has been used in every subsequent wave of the child supplement (which is carried out every two years).

As described in the NLSY-CS handbook (Baker et al., 1993), the Short Form was created in cooperation with Bradley, by choosing about 3 items from each of the Long Form subscales. These items were then collapsed into two subscales for each age group, Cognitive Stimulation and Emotional Support. Items which could clearly be answered by observing the home or the mother and child together for a brief period of time were retained as interviewer response items. Those items which were likely, in the original format, to require some questioning of the mother were translated into a written questionnaire. Interviewers observed the behavior of the mother toward the child and the physical characteristics of the home while they were administering the various instruments required by the NLSY-CS (Baker et al., 1993).

Evidence on the Reliability and Validity of the HOME

As mentioned above, the various versions of the HOME have been used for many purposes by researchers over the years. By examining the results of a sample of these studies, focusing on the available psychometric data on the scale's reliability and validity, areas in which the HOME has been shown to be a well-constructed and valuable measure will become clear. Comparing the Long Form and Short Form of the HOME will also make evident areas in which the Short Form could be improved.

Internal Consistency

Using both alphas and Kuder-Richardson internal consistency coefficients, the subscales of the Long Forms on the whole appear to have substantial internal consistency (see Table 1). Several researchers, however, found low internal consistency for the Variety of Stimulation subscale (Barnard, Bee & Hammond, 1984; Bradley & Caldwell, 1984b; Caldwell & Bradley, 1984; Elardo, Bradley & Caldwell, 1975; Mitchell & Gray, 1981; Sugland et al., 1995), the Modeling of Social Maturity subscale (Bradley et al., 1992; Caldwell & Bradley, 1984; Sugland et al., 1995), and the Acceptance/Lack of Physical Punishment subscale (Mitchell & Gray, 1981; Sugland et al., 1995). These findings are not surprising, since the Variety of Stimulation subscale was originally formed not from factor analysis, as were most of the other subscales, but from miscellaneous items which significantly predicted achievement, and the authors grouped the items together for this reason (Bradley & Caldwell, 1979). Modeling and Acceptance are both shorter subscales (5 items on Modeling, 4 on Acceptance), so their lower alphas may be an artifact of their shorter length. As long as the HOME Long Form is used as a risk measure, the

low internal consistency of these three subscales is not of grave concern.

The Home Short Forms (with different forms for three different age groups), however, have only two subscales. Cognitive Stimulation and Emotional Support, each of which has over 10 items. Data from the NLSY-CS[2] (see Table 2), give alphas for Emotional Support in the range of .35 to .61 for the different age versions, consistently lower than the alphas for Cognitive Stimulation, which range from .50 to .72 (Baker et al., 1993). The low alphas for the Emotional Support subscale likely reflect the lack of a clear underlying construct. In fact, the alphas for the HOME-SF as used in the JOBS Descriptive Study sample (Moore et al., 1995), a mostly-African-American welfare sample from Fulton County, Georgia, show a similar pattern to the NLSY-CS data, but in a noticeably lower range (see Table 3). The alpha for Cognitive Stimulation is .55 and for Emotional Support is .32, based on the Early Childhood version of the HOME-SF.

When the HOME Short Forms are factor analyzed using data from the NLSY-CS, the two subscales are not reproduced, but rather 3 to 7 shorter subscales emerge (Menaghan & Parcel, 1991; Parcel & Menaghan, 1989; Sugland et al., 1995). These shorter subscales, despite having fewer items, have larger alphas than broad subscales of Cognitive Stimulation and Emotional Support (see Table 4). In fact, several researchers using the Short Form either refactored the subscales (Menaghan & Parcel, 1991; Parcel & Menaghan, 1989; Sugland et al., 1995) or created conceptually based subscales (Barratt, 1991) which were smaller and different from the two provided in the NLSY handbook (Baker et al., 1993). This trend is a reflection of the low

[2]The New Chance Evaluation (Quint, Polit, Bos & Cave, 1994) and the JOBS Descriptive Study (Moore et al., 1995) also used the HOME-SF. Data from the New Chance Evaluation indicated low internal consistency of the two major subscales. As a result, four subscales with greater clarity as to underlying constructs were created. In addition, individual items were summarized in terms of three rather than two response options.

internal consistency and lack of clear underlying constructs of interest in the subscales as provided in the current Short Form.

Interrater Reliability

The original designers of the HOME Long Forms set a precedent in their early publications of measuring interrater reliability using percentage of exact agreement between raters as the standard (Bradley & Caldwell, 1979). The HOME manual suggests a level of 90% agreement to ensure adequate interrater reliability (Bradley & Caldwell, 1984). Most researchers have had little trouble attaining this seemingly high level of reliability (although see Barnard, Bee & Hammond, 1984; Bates et al., 1982; and Widmayer et al., 1990, for exceptions). Another measure of interrater reliability, correlation between raters, has been found to range from .66 (Stevenson & Lamb, 1979) to .90 (Parks & Smeriglio, 1986) for the total scale, and from .73 (Parks & Bradley, 1981) to .96 (Allen et al., 1984) for the Maternal Involvement subscale, for example. However, neither percent agreement nor correlation takes into account the likelihood of raters agreeing by chance (Barnard, Bee & Hammond, 1984). Since items on the HOME Long Form are scored on a 0 to 1 scale, the likelihood of two raters agreeing by chance is quite high. Researchers reporting Kappa statistics, which take into account chance agreement and base rates, find Kappas of .80 to .88 (Affleck et al., 1982; Allen et al., 1983; Bradley, Caldwell & Rock, 1988). Overall, these statistics indicate fairly high interrater reliability for the Long Forms.

The Short Forms, on the other hand, have a noticeable lack of information on interrater reliability. These forms were developed for the NLSY-CS in 1986 and have been used in biennial iterations of this survey. In none of the survey rounds has any data been collected on

interrater reliability. This is especially troubling in that the training for survey interviewers is significantly shorter than that typically used for the Long Forms (Baker et al., 1993). Without any data on interrater reliability for the Short Forms, no conclusions can be made about the reliability of the interviewer portions of these scales, and caution should be used in interpreting these data.

Test-Retest Reliability

The HOME Long Forms have been used in longitudinal research many times over the years. Test-retest reliability data are available for various combinations of age groups and time spans, focusing mainly on children ages six months to 5 years. Overall, two generalizations can be made based on a review of the available research. First, test-retest correlations are highest for shorter time spans (averaging about .60 for a one-year span), regardless of the age group being tested (Bradley et al., 1989; Ramey, Yeates & Short, 1984; and Siegel, 1984). Second, when comparing similar time spans, scores become more stable as children get older (Yeates et al., 1983), although data on children over age 5 are limited (but see Bradley et al., 1992, for an example).

As in other realms, test-retest reliability information for the Short Forms is limited to data collected in the NLSY-CS. Here, as with the Long Forms, longer time spans show lower stability, with two-year spans hovering around .50 and four-year spans around .40. Also, the stability of the Cognitive Stimulation subscale is somewhat higher than the Emotional Support subscale (Baker et al., 1993). Overall, the stability of the Short Forms appears only slightly lower than the stability of the Long Forms.

Concurrent and Predictive Validity

Using both concurrent and longitudinal measures, the data show that both Short and Long Form HOME Total scores and many subscale scores are strongly related to various measures of SES such as occupational status (Bates et al., 1982; Bradley & Caldwell, 1984b; Bradley et al., 1992; Bradley, Caldwell & Rock, 1988; Menaghan & Parcel, 1991; Parks & Bradley, 1991), education (Bradley & Caldwell, 1984b; Bradley et al., 1992; Bradley, Caldwell & Rock, 1988; Garrett, Ng'andu & Ferron, 1994; Hollenbeck, 1978; Klebanov, Brooks-Gunn & Duncan, 1994; Menaghan & Parcel, 1991; Parks & Bradley, 1991), income (Bradley & Caldwell, 1984a; Garrett, Ng'andu & Ferron, 1994; Hollenbeck, 1978; Klebanov, Brooks-Gunn & Duncan, 1994) and single parent status (Allen et al., 1984). These results generally show correlations between SES markers and HOME scores of between .3 and .4, although these figures vary between samples. One group of researchers found contradictory results in a sample of infants recruited from intensive care and genetics counseling (Affleck et al., 1982; Allen et al., 1983), but the overwhelming majority of findings indicate higher HOME scores for better educated, wealthier families with two parents.

Additionally, several authors comment on a lack of variability among middle and upper class families due to ceiling effects (Belsky, Garduque & Hrncir, 1983; Mitchell & Gray, 1981; Parks & Smeriglio, 1986; Stevenson & Lamb, 1979; Van Doorninck et al., 1981), although one researcher (Metzl, 1980) demonstrated that the HOME could discriminate between experimental and control groups in a middle class sample. In this study, among normal infants with middle class parents, the experimental groups, which received instruction in methods of increasing

language stimulation in everyday interactions, had higher HOME scores than the control group. While predictions of experimentally induced differences are not often cited as evidence of predictive validity, this evidence is especially important in this case, since it shows that middle class parents have at least some room to improve their HOME scores, despite evidence of ceiling effects.

Both Short and Long Forms have also been shown to predict intelligence and cognitive achievement. Using the Long Form, concurrent correlations between HOME scores and cognitive scores generally range from .3 to .5 (Bradley & Caldwell, 1979; Bradley, Caldwell & Rock, 1988; Kurt, Borkowski & Deshmukh, 1987). Predictive relationships between HOME scores and cognitive scores are stronger as children get older, despite increasing time between assessments. For example, using HOME Long Form subscales at 12 months and Bayley IQ at 12 months, Bradley and colleagues (1989) found predictions between .02 and .28. Prediction from the same 12 month HOME scales to 24 month Bayley was from .17 to .50, and to 36 month Bayley was .20 to .53. This increase in prediction from concurrent to 1 and 2 year longitudinal measures of cognitive development is surprising, but can perhaps be explained by the literature on the stability of measures of intelligence (Pianta & Egeland, 1994). Measures of intelligence at very young ages, such as 12 months, are less stable predictors of later intelligence than measures of intelligence at slightly older ages. It would seem that the HOME, being a more stable measure for younger children, picks up on some aspects of the child's very early environment which are not reflected in intelligence scores until these scores become more reliable.

Two research groups investigated the relationship between cognitive achievement and

HOME scores in different racial groups (Sugland et al., 1995; Berlin et al., 1995). Both reported more or stronger predictions for European-Americans than for African-Americans (Berlin et al., 1995) or than for African-Americans and Hispanics (Sugland et al., 1995). This was true regardless of whether the HOME scores were refactored into race-specific factor subscales (Sugland et al., 1995). The consistency of these findings should not be overstated, however, since the samples were both drawn from the same large study (the Infant Health and Development Program), and data from other samples should be analyzed before drawing any strong conclusions about racial differences in predictive validity.

Data on prediction of intelligence and achievement from the Short Forms, in the NLSY-CS, have been thoroughly investigated. Using bivariate correlations, Baker and colleagues (1993) showed that the Total HOME-SF correlates with PIAT[3] achievement scores about .3 (children ages 5 and up), Digit Span about .2 (ages 7 to 11) and PPVT[4] vocabulary scores about .4 (ages 4 and 10 to 11) when examined concurrently in 1990. The correlations of these cognitive measures with the Cognitive Stimulation subscale are slightly lower, and with the Emotional Support subscale, somewhat lower yet.

Predictive correlations from 1986 HOME-SF Cognitive Stimulation to 1988 and 1990

---

[3] The Peabody Individual Achievement Test, or PIAT is a wide ranging assessment of children's achievement. The correlations reported here refer to Reading Recognition, Mathematics, and Reading Comprehension subscales, as used in the NLSY.

[4] The Peabody Picture Vocabulary Test (Revised), called the PPVT, is a measure of receptive vocabulary designed for children ages two and up (Dunn and Dunn, 1981). Raw scores on the PPVT are converted into standard scores, by month of child's age. While some have criticized this measure for underestimating the cognitive abilities of minority children, it has proven to be a strong predictor of later IQ scores and achievement in both African-American and white children.

PIAT assessments were essentially the same for the three subscales of the PIAT: mathematics, reading recognition and reading comprehension. For children aged 3 to 5 in 1986 (Early Childhood Short Form), correlations averaged .27 in 1988 and .37 in 1990. For children 6 to 9 (Middle Childhood), correlations were .15 in both years. For children 10 and older, correlations were .29 in 1988 and .33 in 1990 (Baker et al., 1993). This pattern of equal or stronger correlations when prediction is for longer periods of time is similar to that found using the Long Forms, and is best documented for the youngest children.

When multiple regression is used to examine the relationship between HOME-SF scores and intelligence or cognitive achievement, percent of variance explained remains about the same as using correlations. Predicting the PIAT and PPVT, figures vary from insignificant to 19% of variance explained by HOME-SF scores, depending on the covariates and the particular subsample of NLSY children used (Barratt, 1991; Desai, Michael & Chase-Lansdale, 1990; Luster & Dubow, 1992; Moore & Snyder, 1991; Sugland et al., 1995). Interestingly, in all three studies which analyzed racial groups separately, predictions of cognitive variables were strongest among Hispanics, and weaker among African-Americans and European-Americans (Luster & Dubow, 1992; Moore & Snyder, 1991; Sugland et al., 1995), which is different from the Long Form findings. As with the Long Form, however, these three studies all use one large data set (the NLSY), so their results should be confirmed with other data sets before strong conclusions are drawn.

Summary of Reliability and Validity Findings

As we have seen, the internal consistency, interrater reliability, test-retest reliability and

concurrent and predictive validity of the HOME Long Forms are impressive. The HOME-SF, on the other hand, is somewhat difficult to evaluate due to the single data source for nearly all of the psychometric information. Based on the available data, the HOME-SF has low internal consistency, considering the large number of items for the two major subscales, and no measure of interrater reliability. On the positive side, the test-retest reliability is comparable to the HOME Long Form, and the concurrent and predictive validity are substantial.

Building from the HOME-Short Form

So far, we have demonstrated the impressive history of research using the HOME Long Forms and the strong, but somewhat less impressive history of the Short Form. At the same time we note several ways in which the HOME-SF could be improved in further survey research.

(1) It seems clear from the varied uses of the HOME-SF over the years that researchers are not simply using it as a risk summary, as it was intended. Many researchers actually want a measure or set of measures which document the full continuum of parenting and home environment, covering as many important aspects as possible. The HOME-SF, with its binary choice format and focus on risk factors rather than on *degrees* of environmental support, was simply not designed for this purpose. A better option for these researchers would be a set of measures which match the HOME-SF in ability to predict development but also measure a greater range of parenting and environment topics and use a full range of responses to capture the full range of variability. This might help address the ceiling effects we have noted.

(2) Concerning the topics covered by the HOME, four main areas of concern have arisen from our review. First, the Short Forms are divided into only two subscales, Cognitive

Stimulation and Emotional Support. Each of these subscales contains items which measure numerous constructs important to development. Neither, however, has a clear underlying meaning when these items are combined. The weaker internal consistency of the two Short Form subscales, as seen in factor analyses and alphas, attests to the extremely varied content of these subscales. Developing a measure with clearly defined subscales and underlying constructs would help both the design and interpretation of research about the home environment.

(3) Some researchers have used portions of the Short Forms or subscales from the Long Forms to investigate the topic of discipline. Due to the focus of the HOME on high risk environments, however, only physical punishment is measured. Discipline is a much more complex topic than simple physical punishment, and the implications of physical punishment are not always clear. As demonstrated by Baumrind and others who have investigated parenting styles, some amount of punishment may be advantageous to development, especially when combined with a caring relationship with an involved parent. One of the discipline questions in the HOME is a good example of this complexity. The question asks what the mother would do if the child hit her. All of her responses are recorded. An effective parent might well both punish the child and discuss the situation with the child. Traditional scoring of the HOME gives any parent who spanks the child in this scenario a bad score, regardless of the other responses. Developing a more complete and clear measure of discipline, and measuring it in a way which allows researchers to identify both problematic and effective discipline would add to the contributions the HOME has made in this area.

(4) As noted above, the HOME is strongly correlated with markers of SES. Several items

on the HOME refer to materials which parents must own in order to receive credit for providing a stimulating environment. One example of this is the number of books the child has of his or her own. Since the HOME has been used extensively in low-income populations, some authors have voiced concerns about this materialistic bias. The developers of the scale intended it to function as a measure of risk, a more accurate measure than the usual income, occupation, and education measures used for SES. Thus, the fact that HOME scores and SES are highly correlated is not at all surprising. Yet the reliance on materials can be a problem, particularly in assessing the sources of cognitive stimulation for children in low income families. As a possible improvement, a new measure could either eliminate these materialistic items, substituting items such as the number of books the child has access to regularly, including library books, books borrowed from friends or relatives, or others; or a new measure could separate materialistic items into a subscale which could be used only when this information is desired.

(5) Further, all HOME scales are based on a combination of mother reports and interviewer ratings. Because the HOME-Short Form was developed for survey use, the partial reliance on interviewers is especially troubling, since no data have been collected on the reliability of survey interviewers using the HOME-SF. Information on interviewer reliability in completing the Short Form is urgently needed.

(6) Finally, when both the Short Forms and Long Forms have been analyzed by racial/ethnic groups, it has been shown that prediction varies in different racial/ethnic groups. A new measure should be piloted on various ethnic groups to ensure its usefulness in all groups.

Table 1
Summary of Selected Current Literature using HOME Long Form[5]

Authors	Type of Sample	N of Sample	Form Used	Internal Consistency	Interrater Reliability	Test-Retest Reliability	Concurrent and Predictive Validity	Comments
Affleck, Allen, McGrade & McQueeney, 1982	High-risk infants recruited from neonatal intensive care or genetics counseling; mainly working class, with some low & middle class.	43	IT		Kappa = .80.		Maternal depression was significantly related to low RE, while high occupation status of the father in father-present homes (32) was related to lower RE & higher AC. Maternal mood, maternal perception of infant temperament & observer-rated infant characteristics together predicted 52% of the variance in TL & similar amounts for each subscale.	Finding of high occupation status of the father in father-present homes related to lower RE & higher AC is contrary to most research.
Allen, Affleck, McGrade & McQueeney, 1983	Oldest infants from Affleck et al., 1982; high-risk, working class.	30	IT		Kappa = .80	From 9 to 18 months, Pearson r = .46 TL, subscales r = .40 to .63, except for RE = .12.	TL & subscales not related to maternal age. Mother education & occupation, & father education were uncorrelated or negatively correlated with TL & subscales. RE & IN had highest negative correlations. Scores on RE, LM, VA, & TL increased from 9 to 18 months.	Negative correlations with SES attributed to special, high-risk, handicapped sample, where handicap may especially stigmatize or frustrate the expectations of middle-class parents.
Allen, Affleck, McGrade & McQueeney, 1984	Both single-mother families & dual-parent families.	22 single-mother, 63 dual-parent	IT		Intraclass correlations of subscales ranged from .86 to .96 (IN).		TL was lower for single mothers at 9 & 18 months & subscales were all in the same direction except for OR. Significant differences were on LM, IN & VA at 9 months; IN & VA at 18 months.	Income was not used as a variable or covariate, but education and occupation were controlled.
Barnard, Bee, & Hammond, 1984	Healthy, low-risk infants of well-educated mothers.	193 total, 161 with complete data	IT	Alpha corrected for number of items: VA = .00 (4 months), OR = .27 (4 & 8), all others including TL range from .30s to .80s.	Range from 86% to 97% agreement; interrater reliability coefficient was as low as .34 (OR).	Developmental generalizability index: .79 for total, .60s for VA, RE, AC, .40s for OR, LM, IN.	4 & 8 month subscales, TL uncorrelated with 12 month Bayley MDI; 4, 8, & 12 month subscales, TL predict 24 month MDI & 48 month Stanford-Binet with correlations as high as .45. This pattern held true for lower and higher education groups and male and female infant groups.	Interrater reliability coefficient compares variability between observers to variability between subjects, so it should be lower than % agreement. However, .34 is noticeably low.
Bates, Olson, Pettit & Bayles, 1982	6-month-old infants recruited from birth notices.	168	IT		88% agreement.		Factor analysis of molecular observational codes with HOME scales & other measures of parents & siblings, found that IN & LM loaded with dependency & occupation level, while VA, OR, & RE loaded with Maternal support, satisfaction, & high desirability--low aggression. AC loaded with mother age, structure, & nurturance.	

[5] A key to subscale abbreviations can be found on the last page of Table 1. Forms used are abbreviated as IT for Infant/Toddler, EC for Early Childhood, MC for Middle Childhood, and PA for Pre-Adolescent.

Table 1 (Continued) Summary of Selected Current Literature using HOME Long Form

Authors	Type of Sample	N of Sample	Form Used	Internal Consistency	Interrater Reliability	Test-Retest Reliability	Concurrent and Predictive Validity	Comments
Belsky, Garduque & Hrncir, 1983	Small sample of middle- & working-class infants	64: 22 (12 months), 20 (15), & 22 (18)	IT		Correlations were over .80 on all subscales.		Free-play behavior (performance) related more consistently to the 6 subscales than executive capacity, which was more related to the HOME than elicited play (competence).	Authors note a general lack of significance in their findings may be due to the insensitivity of the HOME to variation within generally supportive environments.
Berlin, Brooks-Gunn, Spiker & Zaslow, 1995	IHDP: Low-birth weight, premature infants selected at birth & followed to age 3. Intervention group of 377 & control group of 608.	486, Control group only	EC	Learning (LM, LA, AST, VA), Whites, .82. Blacks, .87; RE, Whites, .61, Blacks, .66.	Over 90% agreement.		Concurrent White: RE (R^2=.05) & Learning (.12) predict receptive language ability. RE (.04) predicts behavior problems. Black: Learning (R^2=.14) predicts receptive language ability.	Authors comment that many Learning items require material resources, which may be unavailable in lower SES, & may overlook effective parenting possibilities for poor mothers.
Bradley & Caldwell, 1979	Volunteer families from Little Rock, Arkansas area; lower to middle class.	242	EC	KR-20 from .53 to .83 for all individual scales. TL, .93	95% agreement.	From 3 years to 4.5 years scales ranged from .05 to .70	Concurrent: 3 year scales to IQ all significant, range from r=.25 (PE) to .55 (TL); 4.5 year scales to IQ, LM, LA, AST, VA, & TL significant, range from .08 (AC) to .58 (TL). Predictive: 3 year scales to 4.5 year IQ all significant, range from r=.17 (MO) to .54 (TL).	Shortened EC Form from 80 items to 55 items. Scale designed primarily as a screening instrument.
Bradley & Caldwell, 1980	Volunteer sample from Little Rock	72	IT			From 6 to 12 months scales ranged from .5 to .6, except for VA (not reported).	Concurrent: 6 & 12 month scales to 12 month Bayley MDI, goal directedness & language use were inconsistent. Predictive: Multiple R of all scales at 6 & 12 months predicting 3 year IQ was .80 for boys & .76 for girls. Best predictors were LM (6 & 12 months) for boys & LM (12), IN (12) & RE (12) for girls.	12 month scales more predictive than 6 month, although highly correlated.
Bradley & Caldwell, 1984a	Little Rock sample, drawn from Caldwell, Elardo & Elardo (1972).	79, intact families only	IT	KR-20 from .44 to .68.			12 & 24 month subscales in intact families by gender, race (Black v White) & SES. Most consistent effects were for crowding & birth order. Especially related to status variables were OR, LM & IN.	24 month scales predicted better than 12 month scales.
Bradley & Caldwell, 1984b	Longitudinal Observation & Intervention Study sample	174	IT	From .89 (OR) to .44 (VA), with all others in .60s & .70s. TL, .89.	Criteria of 90% agreement.	From 6 to 12 months, range from .29 to .62. From 6 to 24 months, range from .27 to .64. From 12 to 24 months, range from .30 to .77.	All subscales & TL are correlated with mother & father education, father presence, father occupation, & crowding at 6, 12, & 24 months, although 6 month correlations were weaker. 6, 12, & 24 month subscales predict 3 year IQ & 4.5 year IQ at correlations of .20s to .60s, with very little change in prediction between 3 & 4.5 years. Multiple correlation between 6 month TL & language development at 37 months measured by Illinois Test of Psycholinguistic Abilities is .41, & 24 month TL with ITPA is .65.	Training included accompanying experienced interviewer on at least 10 visits.

Table 1 (Continued) Summary of Selected Current Literature using HOME Long Form

Authors	Type of Sample	N of Sample	Form Used	Internal Consistency	Interrater Reliability	Test-Retest Reliability	Concurrent and Predictive Validity	Comments
Bradley, Caldwell, Brisby, Magee & Whiteside, 1992.	10 to 15 year old children with disabilities; middle-, working- & lower-class families.	117	MC, PA	For PA only, PH, .71; LM, .66; MO, .53; Instructional & regulatory, .71; VA, .61; Acceptance & responsivity, .75; TL, .88.	94%	40 children had MC data from 18 months earlier. TL=.71, PH= .57, LM=.57. Cross subscale r ranged from .02 to .59.	Correlations of PA scales with SES, parent IQ, & parent social skills are highly concentrated between .30s & .50s. Correlations with child competence were nonsignificant for PH, MO, & Acceptance & responsivity. Others subscales had a mix of nonsignificant & significant, but small, ranging from .03 to .31, except for Instructional & regulatory, which ranged from .23 to .44.	Preliminary, 80 item version of Preadolescent HOME. Testing & item reduction continues for the PA.
Bradley, Caldwell & Rock, 1988	Normal children, followed from Elardo, Bradley & Caldwell (1975), to 4th & 5th grade.	42	IT, MC	Alphas for MC from .52 to .80, .67 median, & .90 for TL.	93% agreement, kappa=.88.	RE, LM & IN were most stable.	Correlations of MC with SES indicators ranged from .2 to .5, median of .43. Correlations with academic achievement had median of .33, lower than IT. 6 month IT did not predict 10 year achievement. 24 month IT predicted achievement in language arts, & 10 year MC predicted language arts & mathematics.	Stability of IN was important to achievement at age 10.
Bradley, Caldwell, Rock, Barnard, Gray, Hammond, Mitchell, Siegel, Ramey, Gottfried, & Johnson, 1989	6 sites across the US, various sampling techniques	931	IT, EC			From 6 to 12 months, TL, .59; from 24 to 36, .73	For IQ, prediction from 12 month HOME subscales increase between 12 (.02 to .28) & 24 months (.17 to .50), then remained stable at 36 months (.20 to .53). Similar for 24 month HOME. Best predictors of IQ were play materials & involvement.	When both Bayley & TL were below -1 SD, half of the children were also below -1 SD on 3 year IQ & 95% were below the mean. At 24 & 36 months, about 90% of those who scored 1 standard deviation above or below the mean TL score were above or below 100 on 3 year IQ.
Bradley, Mundfrom, Whiteside, Casey & Barrett, 1994	IHDP	870, 299 White, 477 Black, 94 Hispanic	IT, EC		90% agreement criterion.		Among Blacks & Whites, both the IT & EC subscales were essentially confirmed by exploratory factor analysis, while Hispanics had different factor structures. The Hispanic group had more factors, which accounted for less of the variance.	
Bradley, Whiteside, Mundfrom, Casey, Kelleher, & Pope, 1994	IHDP	243, portion of control group in poverty only	IT, EC	Total of RE, AC/AC, LM/LM, & VA, plus one item on safety. 12 months: Whites .79, Blacks .82, Hispanics .74; 36 months: Whites .87, Blacks .86, Hispanics .84.			Using cutoffs on 6 protective variables (AC/AC, LM/LM, VA, RE, safe play area, & crowding), prediction of non-resilience was very high (98% & 94%) while prediction of resilience was poor (15% & 20%).	LM/ LM, AC/AC, VA/VA, RE/RE are summed between IT & EC Forms.
Bradley, Whiteside, Mundfrom, Casey, Caldwell, Barrett, 1994	IHDP	of 985 875 IT 819 EC	IT,EC	IT-.89 range from .44 to .89 EC-.93 range form .53 to .88	90% agreement criterion		IT revealed significant main effects for site, maternal education, and ehtnicity. Significant intervention effects were noted on the Learning Materials subscale, Learning stimulation, Modeling, Variety, and Acceptance subscales.	Factor analyses of the HOME Inventory revealed that intervention and follow-up groups had similar underlying structures

Table 1 (Continued) Summary of Selected Current Literature using HOME Long Form

Authors	Type of Sample	N of Sample	Form Used	Internal Consistency	Interrater Reliability	Test-Retest Reliability	Concurrent and Predictive Validity	Comments
Bradley, Mundfrom, Whiteside, Caldwell, Casey, Kirby, Hansen, 1994	IHDP/Low birth weight preterm infants	included 591 of 985 African American, white, Hispanic mothers and their infants	IT-12 months of age EC-36 months of age		90% agreement criterion		Ethnicity, income, maternal education, and marital status, maternal age, gestational age, were used as predictor variables. After controlling for other variables the portion of variance in HOME scores explained by income was low (IT-5.1%, EC-4.2%). In both IT and EC HOME scores ethnicity contributed the largest portion of explained variance.	Correlations between HOME scores and four child characteristics measured at 36 months indicated that the quality of the home environment as measured by the HOME Inventory, is related to children's development.
Caldwell & Bradley, 1984	Standardization sample, not described.	On IT, 174 used for KR-20, 91 for test-retest. No N reported for EC.	IT, EC	KR-20 for IT: RE, .72; AC, .67; OR, .89; LM, .77; IN, .69; VA, .44; TL, .89. For EC: LM, .88; LA, .65; PH, .83; RE, .75; AST, .60; MO, .53; VA, .69; AC, .59; TL, .93.		6 to 12, 6 to 24 & 12 to 24 months, VA & TL, highest Pearson & intra-class correlations, .50s to .70s. Lowest was AC, .23 to .32. 3 to 4.5 years, highest were LM & TL, .70. Lowest, AC, .05. Others .21 to .48.		This is the HOME manual for IT & EC forms, with early data on MC form. No information is given about the samples, although means & standard deviations are provided. Authors note, "We confidently expect that in future analyses of the items we will find that some items do not correlate significantly with the child's development."
Chipeur & Plomin, 1992	Colorado Adoption Project sample, adopted & non-adopted sibling pairs.	92 pairs.	IT			Over a two-week interval at 12 months, 3 of 45 items had non-significant stability. 12 to 24 months, non-shared scale r=.64, shared r=.77.	12 month Shared & Non-Shared correlated with 12 month Bayley (.16, .20). At 24 months, both correlated with 24 month Bayley (.38, .16).	One item, expressiveness of mother, was dropped due to no variability in this middle-class sample. Quantitative scoring rather than dichotomous was used, due to greater stability from 12 to 24 months.
Coll, Hoffman & Oh, 1987	Primiparous Caucasian low- & middle-class mothers, half adolescent, half adult, with 4-month-old infants	50	IT		89 to 94% agreement.		When SES was controlled, adolescent mothers had lower RE, IN, & TL. TL differed by one standard deviation (31 versus 35) between adolescent & adult mothers. With SES not controlled, Adolescent mothers had lower AC & VA scores.	
Gottfried & Gottfried, 1984	Middle-class sample.	130, plus 40 pilot	IT, EC		91 to 100%, mean of 97% agreement.		TL at 15 months & 39 months predicted Bayley MDI at 12, 18, & 24 months, & McCarthy at 30, 36, & 42 months.	Special pilot sample used to train interviewers.
Hollenbeck, 1978	Caucasian sample, distributed around middle-class.	70	IT		Intraclass correlation = .81		Family income, maternal & paternal education correlated with all but OR & LM.	LM mean was 7.1 out of a possible 8, indicating a ceiling effect.

Table 1 (Continued) Summary of Selected Current Literature using HOME Long Form

Authors	Type of Sample	N of Sample	Form Used	Internal Consistency	Interrater Reliability	Test-Retest Reliability	Concurrent and Predictive Validity	Comments
Johnson, Breckenridge, & McGowan, 1984	Mexican-American sample, includes some siblings.	118 control, 96 program.	EC or IT?		Over 90% agreement.		Comparing HOME scores to observational data from the Mother-Child Interaction, which is a semi-structured, task-oriented session rated on 7 broad dimensions of the interaction, correlations between HOME & MCI were few & small. IQ at age 3 was correlated with lower scores on AC, & higher IN.	The relationship between AC & IQ is negative, which is the opposite of most samples.
Klebanov, Brooks-Gunn & Duncan, 1984	IHDP Sample, only those whose data could be matched to census.	895	EC	Learning (LM, LA, AST, VA), .87; PH, .74; WA, .64.	At least 90% agreement.		Family poverty variables correlated with HOME scores. Having more families with incomes less than $10,000 in the neighborhood was related to less RE & poorer PH.	Training of interviewers was 2 days, with criterion of 90%, plus periodic checks.
Kurtz, Borkowski & Deshmukh, 1987	From Nagpur, India (large city), selected from school entrance exams to represent the full range, with over sampling at the top & bottom of the distribution.	60, 30 1st grade, 30 3rd grade	MC		.93, unspecified index.		SES & TL correlated .81, with TL a better predictor of concurrent memory & achievement. TL was significantly related to achievement & marginally related to recall, while child & mother IQ were not significant in either multiple regression equation. TL was moderately correlated with child IQ (.38 & .40) & moderately to highly with mother IQ (.44 & .61).	The MC scale was translated into Marathi language.
Lozoff, Park Radan, Wolf, 1995	Cost Rican Infants from urban Hatillo /homogeneous community inhabited by lower-middle class residents Little Rock, Arkansas infants from lower-middle class families	191 total, 161 with complete data 174	IT	KR-Total: .89 Little Rock; .84 Costa Rica KR was lower on Subscale 3, Organization of the Environment for Costa Rican sample .31 vs .89			Both concurrent and predictive validity of the HOME in respect to a child's cognitive function was compared in Cost Rica and US by correlations between the HOME subscale and total scores of intellectual assessments. There were no significant correlations between Costa Rican HOME subscale or total scores and Bayley MDI. The correlations between HOME scores in infancy and cognitive function at 4 ½ or 5 years were statistically significant in both samples, but significantly lower in the Costa Rican sample than in the US sample.	The psychometric properties of the HOME in the Costa Rican children were similar to those of the Little Rock sample
Luster, Perlsadt, McKinney, Sims, Juang, 1996	Low income families, expecting first child, and had not completed highs school Enrolled in Family Ties Program	83	IT				Teens who provided more relatively supportive environments: were less depressed and more empathic, had infants were heavier at birth and less irritable at 12months, received more support from the father of the baby, and lived in safer neighborhoods.	
Matheny, Wilson & Thoben, 1987	Twins from a broad range of SES families were followed from 12 to 24 months	112 (56 pairs)	IT		.86, unspecified index.		None of the scales correlated with a measure of temperament at 12, 18 or 24 months.	

Table 1 (Continued) Summary of Selected Current Literature using HOME Long Form

Authors	Type of Sample	N of Sample	Form Used	Internal Consistency	Interrater Reliability	Test-Retest Reliability	Concurrent and Predictive Validity	Comments
Metzl, 1980	Normal firstborns with married, middle-class parents measured at 6 weeks & 6 months.	60	IT		94% agreement.		Treatment groups improved TL more than the control, which also increased.	This showed differentiation between experimentals & controls by total HOME scores in middle-class families.
Mitchell & Gray, 1981	Firstborn infants *recruited from prepaid* medical group. Highly educated mothers.	144	IT	TL, .77 to .86; RE, .65 to .68; AC, .38 to .63; OR, .25 to .39; LM, .59 to .74; IN, .58 to .66; VA, .00 to .37. Developmental generalizability index, based on repeated measures ANOVA of items within scales, TL= .74, subscales= .23 to .68.		Based on repeated measures ANOVA, at 4, 8, 12 & 24 months, stability of TL = .79, subscales = .44 to .66. Lower in high education subsample.		High education subsample had higher scores & lower variance than low education, likely due to ceiling effect. 24 month rating modified for laboratory, & not reported here.
Mundfrom, Bradley, Whiteside, 1993	IHDP	985	IT, EC				Factor structures of the IT and EC HOME provide evidence that these two forms of the Inventory measure what they were designed to measure. Four of the factors in the IT analysis and all five in the EC correspond directly to existing subscales on these two versions of the Inventory.	Characteristics of the home environment were measured at two different points in the children's development -12 and 36 months.
Parks & Bradley, 1991	Infants recruited from birth records or a well-baby clinic, half from suburban & half from urban; wide range of SES; sample divided into low & middle/high SES subsamples.	155, 55 low SES, 87 mid/ high SES	IT	TY, .65; IN, .71.	r=.93 LM & r=.73 IN (Pearson)		Both scales were higher in higher SES. For the full sample, the main effect of LM & IN together, when entered first, significantly predicted locomotor, eye-hand, & general development, as did the interaction of the two. Neither was significant when entered second. *Only* the interaction was important in predicting personal-social & hearing-speech development, such that only when LM was high did IN have a significant effect on both domains. A similar pattern was found in middle/high SES (n=55) families for general development, but no main or interaction effects were significant for low SES.	
Parks & Smeriglio, 1986	Mothers 18 years & older recruited from a well-baby clinic or birth records. Sample was divided into low, middle, & high SES subsamples.	126	IT	TL, .77 (low SES), .69, (middle), .70 (high).	Pearson r=.90 full sample.		Knowledge of the effects of parenting on infant development correlated with TL in low but not middle or high SES samples.	Both TL & parenting knowledge scales showed ceiling effects & reduced standard deviations in middle & high SES groups.

Table 1 (Continued) Summary of Selected Current Literature using HOME Long Form

Authors	Type of Sample	N of Sample	Form Used	Internal Consistency	Interrater Reliability	Test-Retest Reliability	Concurrent and Predictive Validity	Comments
Ramey, Yates, & Short, 1984	A high-risk mostly Black sample from North Carolina; one experimental day-care group & one control group followed from birth to school entry.	54 experimental, 52 control.	IT, EC		Over 90% agreement.	TL 6 to 42 months, .24 control, .27 experimental. Median one-year TL stability .48 experimental & .62 control.	Total HOME Index prediction to IQ was best for 48 month IQ. Both the predictive value of the HOME Index & of maternal IQ increased with age, interpreted by the authors as an increase in predictability of IQ. There were no interactions of maternal IQ or HOME Index with group membership, meaning that the HOME Index changed similarly in experimental & control groups.	A total HOME Index was computed by standardizing & summing TL over the 4 occasions of measurement.
Rock, Head, Bradley, Whiteside, Brisby, 1994	Non institutionalized visually impaired children 6 months to 6 years Recruited from a larger study of the home environment of children with disabilities	31	IT, EC				Correlations between the HOME scores, scores on measures of family ecology, and measures of children's competence were generally in line with expectations. Families of children with visual impairments scored about the same on standard forms of the IT and EC as did families in the norm group for the intervention. Correlation patterns between the HOME and measures of support varied depending on the child's age and the area of support examined. Correlations between IT scores and measures of children's competence only revealed significant correlations between the Acceptance of Child subscale and children's competence. Some correlations are interpreted cautiously due to the small sample size	Constructed modified versions of the IT and EC (containing some new items and some adapted scoring procedures) to see if more useful info would emerge when these versions were applied to families of visually impaired children. According to the authors, for predictive purposes there seems to be no advantage to using the alternate form of the HOME because the order of scores is essentially the same.
Siegel, 1984	Pre-term & full-term samples, each in 2 cohorts.	121 full-term, 117 pre-term	IT, EC			12 months to 36, range from .33 to .78. From 12 to 60, .39 to .41. From 36 to 60, .40 to .71.	3 year TL predicted 5 year McCarthy ability test, but 12 month TL did not predict 5 year McCarthy.	
Stevenson & Lamb, 1979	Infants recruited from birth announcements, upper-middle class.	40	IT		Pearson r=.66, Spearman-Brown =.80.		TL correlated with initial infant sociability at 12 months, mainly due to RE, which essentially measures maternal sociability. TL did not correlate with infant sociability during testing, Bayley, or psychological development.	Possible problem using HOME with middle-class sample noted.
Sugland, Zaslow, Smith, Brooks-Gunn, Coates, Blumenthal, Moore, Griffin & Bradley, 1995	IHDP, all children with 36 month data; ethnically diverse with low birth weight infants.	819	EC	All races together, TL, .89; AC, .54; RE, .65; PH, .74; MO, .36; LM, .80; LA, .57; AST, .53; VA, .49. Black, range from .18 (MO) to .73 (LM). Hispanic from .04 (VA) to .76 (LM). White from .30 (MO) to .77 (LM).	Over 90% agreement.		Learning, RE & PH were significant predictors of 3 year Stanford Binet & CBCL in all 3 racial groups, strongest in Whites.	Only subscales with alpha > .60 in full sample, plus Learning (LM, LA, AST, VA) composite were included in concurrent validity analyses.

Table 1 (Continued) Summary of Selected Current Literature using HOME Long Form

Authors	Type of Sample	N of Sample	Form Used	Internal Consistency	Interrater Reliability	Test-Retest Reliability	Concurrent and Predictive Validity	Comments
Van Doorninck, Caldwell, Wright & Frankenburg, 1981	Sample of low-income & middle-income children were followed from 12 months to elementary school (grades 1 to 6).	50 low-income, 21 middle-income.	IT			Comparing siblings tested at the same age, intervals from 10 to 30 months, r=.86 (N of pairs = 36).	HOME & elementary school achievement problems were not significantly correlated in middle-income sample, due to restricted variability in HOME scores. IN, VA & TL correlated with achievement among low-income, although the effect was actually a threshold of 30.5 TL which related to achievement. Low TL predicted 14 of 20 children's low achievement scores & high total predicted 20 of 30 high scores.	Lack of prediction in middle-class sample using correlations, due to restricted variability.
Watson, Kirby, Kelleher, Bradley, 1996	IHDP	of 985 608 control group infants selected	EC				Total HOME scores differed significantly between the groups with the total score for the poor group being approximately 1 standard deviation lower than that for the nonpoor group. Scores for all but one HOME subscale (Acceptance) were significantly lower for the poor sample. Normative data for the HOME Inventory (Caldwell & Bradley, 1984) used for comparative purposes showed similar scores. Regression model including poverty, race, site and representative environmental, maternal and child variables accounted for 60% of variance in total HOME scores. As expected, there was a higher percentage of African American families (73%) in the poor as compared to norpoor (44% group).	Site was included as an environmental variable. Poverty thresholds were based on household income. Families were considered to be living in poverty if they were below thresholds at the 12, 24, and 36 month assessemnts.
Weintraub & Palti, 1991	Intervention group: Five year old children exposed to an early intervention program from birth to two years Control group: Sample from existing data of 900 children attending health centers in Jerusalem	study population 36 matched controls 36	EC				Intervention group scored somewhat higher on HOME than control group but the difference was not significant No statistically significant difference in the HOME score between controls or exposed groups Difference in HOME score was negligible between the upper and lower maternal education sub-groups in the exposed group	Lack of statistical difference is attributed to a general increase in parental knowledge about enriching the home environment/the contribution of the home to the development of their children
Widmayer, Peterson, Lamer, Carnahan, Calderon, Wingerd & Marshall, 1990	Infants recruited from Haitian-American neighborhoods in Florida, some from rural & some from relatively better off urban.	66, 22 rural, 44 urban.	IT		79% agreement		The rural sample had higher AC & lower LM, but TL scores did not differ. TL was not related to birth weight, crowding or psychomotor development, but was related to Bayley MDI.	

Table 1 (Continued) Summary of Selected Current Literature using HOME Long Form

Authors	Type of Sample	N of Sample	Form Used	Internal Consistency	Interrater Reliability	Test-Retest Reliability	Concurrent and Predictive Validity	Comments
Yeates, MacPhee, Campbell & Ramey, 1983	High-risk, referred sample of Black infants, followed from 6 to 42 months	46	IT, EC (80 items)		Over 90% agreement	Stability increased with age: 6 to 18 = .43, 18 to 30 = .52, 30 to 42 = .68.	Hierarchical regression & path analysis show that, when maternal IQ is considered, THI does not predict child IQ at 24 months, predicts only in combination with maternal IQ at 36 months, & predicts significantly in addition to maternal IQ at 48 months.	THI was created from summing & standardizing all TL scores prior to measurement of an outcome.
Zaslow, Berlin, Brooks-Gunn, Coiro, Spiker, Moore, Blumenthal & Brown, 1995.	IHDP sample, control group.	486, 204 Whites, 282 Blacks.	EC	Learning (LM, LA, AST, VA), .82 White, .87 Black; WA, .61 White, .66 Black.	Over 90% agreement.		Learning correlated with quality of assistance .37 White, .25 Black. RE correlated with supportive presence .27 White, .22 Black. Learning predicted PPVT-R & CBCL, RE predicted PPVT-R & CBCL in Whites only.	

Subscale Names and Abbreviations
Adapted from Caldwell & Bradley, 1994

Infant/ Toddler (IT)	Early Childhood (EC)	Middle Childhood (MC)	Name	Abbreviation
Factor Number of Subscale				
1	4	1	Responsivity	RE
2	8	3	Acceptance (avoidance of restriction and punishment)	AC
3			Organization (of physical and temporal environment)	OR
4	1	4	Learning Materials	LM
5			Involvement	IN
6	7		Variety (of stimulation)	VA
	2		Language Stimulation	LA
	3	7	Physical Environment	PH
	5		Academic Stimulation	AST
	6		Modeling (of social maturity)	MO
		2	Encouragement of Maturity	EM
		5	Family Participation in Enrichment	FAM
		6	Family Integration	FI
			Total	TL

Table 2
Summary of Selected Current Literature using HOME Short Form[6]

Authors	Type of Sample	N of Sample	Form Used	Internal Consistency		Interrater Reliability	Test-Retest Reliability	Concurrent and Predictive Validity	Comments
				Cognitive	Emotional				
Baker, Keck, Mott & Quinlan, 1993	NLSY 1986. Firstborns, poor, minorities & children born to younger mothers are over-represented.		IT,* EC, MC 6 to 9, MC/ 10+				Total 1986 to 1988=.54; 1986 to 1990=.45. CS 1986 to 1988=.46; 1986 to 1990=.42. ES 1986 to 1988=.38; 1986 to 1990=.28.	Predictive correlations for all children age 3 & up from CS 1986 to PIAT 1988 & 1990: math, .24 (88), .26 (90); reading recognition, .22 (88), .30 (90); reading comprehension, .22 (88), .28 (90).	Interviewers were trained for 2.5 days, including all *instruments for the* NLSY. The first interview for each interviewer was taped & reviewed.
	NLSY, 1988.		IT, EC, MC 6 to 9, MC/ 10+						
	NLSY, 1990.	974, IT; 1124, EC; 1531, MC; 1132, MC/10+.	IT, EC, MC 6 to 9, MC/ 10+	IT, .50 EC, .72 MC, .67 MC/10+, .62	IT, .35 EC, .54 MC, .61 MC/ 10+, .58			Concurrent correlations for the entire sample of TL: PIAT, .29 to .31; Digit Span, .21; PPVT, .37 to .47.	Some data was not useable (11% of children under 3) because interviewers did not have enough time to witness *interactions between* mothers & children under age 3.
Barratt, 1991	NLSY 1986, firstborn children ages 6 to 7 born to mothers ages 14 to 16.	258	MC	Conceptually created scales: Reading & Enrichment =.59, 9 items, Responsibility =.67, 5 items, Observed Involvement =.71, 5 items, & Lack of Punitiveness =.64, 3 items.				In multiple regression, the HOME scales were related to PPVT scores (R^2=.07) but not PIAT reading or math achievement, when SES and mother IQ were included.	Observed Involvement was observational, all other scales were mother report only.
Desai, Michael & Chase-Lansdale, 1990	NLSY 1986, all children ages 3 to 5.11.	1653	EC	.69	.49			For both boys & girls, each yes item on CS related to about 2 points on the PPVT. In multiple regression on PPVT, using numerous SES and status variables as controls, CS & SE together for girls had R^2=.05, for boys R^2=.08. CS & SE were positively related to mothers' work outside the home, but mainly in low income families.	They noted that the Short Form is likely to distinguish between adequate & poor home environments, but seems less useful at the upper end of the spectrum. *Two resources reflected in* CS, Financial & Intellectual engagement.

[6] Forms used are abbreviated as IT for Infant/Toddler, EC for Early Childhood, MC for Middle Childhood,and MC/10+ for Middle Childhood form altered for adolescents. Subscales are abbreviated as CS for Cognitive Stimulation and ES for Emotional Support.

Table 2 (Continued) Summary of Selected Current Literature using HOME Short Form

Authors	Type of Sample	N of Sample	Form Used	Internal Consistency: Cognitive / Emotional	Interrater Reliability	Test-Retest Reliability	Concurrent and Predictive Validity	Comments
Hannan & Luster, 1991	NLSY 1986, all children 12 to 23 months old.	602	IT	TL, .56.			Using age at first birth, education, self-esteem, AFQT, *income, presence of partner, & number of children* as risk indicators, the likelihood of as a bottom tercile TL increased with each risk factor. Children with all 6 had 88% likelihood of low TL while children with 0 had 11% likelihood.	HOME scores used as *dependent variable*.
Luster & *Dubow, 1992*	NLSY 1986, all *children 3 to 8.*	1336, *ages 3* to 5; 832, ages 6 to 8.	EC, MC	TL, .70 for both EC & MC.			Mother's AFQT intelligence score & TL predicted similar amounts of variance in PPVT IQS of preschoolers, although betas were smaller in Hispanic & Black. For children 6 to 8, maternal intelligence was about twice as important as TL in predicting PPVT (both concurrent), *with smaller betas for Blacks & Whites.*	Effects were found not to vary by year of age, but rather to jump between preschool & elementary school ages.
Menaghan & Parcel, 1991	NLSY 1986, children 3 to 6 with employed mothers.	708	EC, MC	3 factor-based scales: cognitive stimulation =.69, EC, .59, MC; emotional warmth-support =.69, .80; & physical environment =.60, .55. *Combined* =.71, EC & MC.			Using multiple regression, Mexican-Americans & Blacks had lower TL than Whites. Internal locus of control, high self-esteem, more schooling, fewer children & occupational complexity relate to higher TL. Maternal AFQT not significant when schooling is considered.	Factor-based scales based on Parcel & Menaghan, 1989, did not include all items from HOME SF.
Moore & Snyder, 1991	NLSY 1986, all firstborn children ages 3 to 7.	626, White; 376, Black; 240, Hispanic	EC, MC				Whites' (.04), Blacks' (.02), & Hispanics' (.08) TL added significantly to the variance explained in PPVT using multiple regression.	Authors rescaled items using the full category answers rather than *dichotomous*. The new scores correlated .90 with the original scores, & did not affect analyses.
Parcel & Menaghan, 1989	NLSY 1986, all children.	1505, 0 to 2 years; 1391, 3 to 5; 1218, 6 & up.	IT, EC, MC all ages	IT: Stimulation=.72, 7 items; *Warm=.71, 5; Non-Punitive* =.50, 3. EC: Cognitive Stimulation=.72, 7; Accept anger=.77, 2; Warm =.69, 4; Physical Environment=.60, 4; No Violence=.59, 2. MC: Paternal=.87, 3; Self Care =.76, 5; Warm=.80, 4; Opportunities=.59, 7; Physical Environment =.55, 5.			Concurrent correlations were found between many of *the factors for all 3 age groups & measures of family* background & child characteristics.	Not all items loaded on any *factors*. Each of the 3 age-groups showed similar factors of warmth & cognitive stimulation.
Menaghan & Parcel, 1991	NLSY 1986, children 3 to 6 with employed mothers.	708	EC, MC	*3 factor-based scales:* cognitive stimulation =.69, EC, .59, MC; emotional warmth-support =.69, .80; & physical environment = .60, .55. Combined =.71, EC & MC.			*Using multiple regression, Mexican-Americans & Blacks* had lower TL than Whites. Internal locus of control, high self-esteem, more schooling, fewer children & occupational complexity relate to higher TL. Maternal AFQT not significant when schooling is considered.	*Factor-based scales based* on Parcel & Menaghan, 1989, did not include all items from HOME SF.

Table 2 (Continued) Summary of Selected Current Literature using HOME Short Form

Authors	Type of Sample	N of Sample	Form Used	Internal Consistency		Interrater Reliability	Test-Retest Reliability	Concurrent and Predictive Validity	Comments
				Cognitive	Emotional				
Sugland, Zaslow, Smith, Brooks-Gunn, et al, 1995	NLSY 1988, all children ages 3 to 5.	1541	EC	All races=.71 Hispanic=.69 Black=.72 White=.70	All races=.59 Hispanic=.60 Black=.52 White=.57			Concurrent multiple regressions showed that all-race subscales predicted PIAT scores more consistently for Whites, although R^2 was similar in Whites (.06) & Hispanics (.07), and smaller for Blacks (.03). Only observed punishment was related to lower PIAT in Hispanics, but it related to higher PIAT in Blacks. Results predicting BPI were similar, but without differences in R^2 between races. Using race-specific factor scales did not increase prediction. Factor analysis was performed in the full, all-race sample & in each racial group separately. Factors were quite similar between races, with some variations in the order of factor emergence (importance).	All-race factors were stimulation, school preparation, involvement, physical environment, & observed punishment.
				All-race factor scales ranged from .51 to .86, & separate-race factor scales from .53 to .85, with no noticeable pattern of differences between races.					

Table 3
Factor Analysis and Internal Consistency of HOME Scales
From JOBS data, N=790

Scale	Number of Factors	Alpha	Number of Items
Total HOME	7	.56	26
Cognitive Stimulation	3	.55	14
Emotional Support	4	.32	12

Table 4
Factor Loadings and Internal Consistency of Factor-Based Scales
From JOBS data, N=790

Factor Scale	Alpha	Number of Items	Items in Scale	Factor Loadings
Cognitive F1: Teaching at home	.80	4	Numbers Alphabet Colors Shapes	.89 .89 .85 .69
Cognitive F2: Physical Environment	.62	4	Clean Not Cluttered Safe Not Dark or Monotonous	.83 .79 .62 .48
Cognitive F3: Cognitive Enrichment	.52	6	Child Owns 10+ Books Child Has 5+ Tapes/Records Family Gets 1+ Magazines Regularly Trips to Museum Mother Reads to Child Child Goes on Outings	.60 .59 .56 .56 .47 .42
Emotional F1: Warmth & Sociability	.42	4	Mother Talks to Child Mother Shows Affection Positive Tone of Voice to Child Introduces Child to Interviewer	.67 .65 .65 .46
Emotional F2: Father Contact	.56	2	Child Sees Father Daily Child Eats with Both Parents Daily	.83 .83
Emotional F3: Observed Punitiveness	.53	2	Mother Physically Restricted Child Mother Slapped Child	.80 .78
Emotional F4: Hit Back, Spanking, TV	.29	3	Spanked 0 or 1 Times in Last Week Mother Would Not Hit Child Back Hours TV on per Week	.72 .62 .58

References

Affleck, G., Allen, D, McGrade, B.J., & McQueeney, M. (1982). Home environments of developmentally disabled infants as a function of parent and infant characteristics. American Journal of Mental Deficiency, 86(5), 445-452.

Allen, D.A., Affleck G., McGrade, B.J., & McQueeney, M. (1984). Effects of single-parent status on mothers and their high-risk infants. Infant Behavior and Development, 7, 347-359.

Allen, D.A., Affleck, G., McGrade, B.J., & McQueeney, M. (1983). Characteristics of the Home Observation for Measurement of the Environment inventory in a sample of high-risk/developmentally disabled infants. Infant Behavior and Development, 6, 53-60

Baker, P.C., Keck, C.K., Mott, F.L. & Quinlan, S.V. (1993). The NLSY child handbook, revised edition: A guide to the 1986-1990 National Longitudinal Survey of Youth child data, Columbus, OH: Center for Human Resource Research.

Barnard, K.E., Bee, H.L. & Hammond, M.A. (1984). Home Environment and cognitive development in a healthy, Low-risk sample: The Seattle study. In A.W. Gottfried (Ed.), Home environment and early cognitive development: Longitudinal research, (pp. 117-149) Orlando, FL: Academic Press, Inc.

Barratt, M.S. (1991). School-age offspring of adolescent mothers: Environments and outcomes. Family Relations, 40, 442-447.

Barrera, M., Rosenbaum, P. & Cunningham, C. (1986). Early home intervention with low-birth-weight infants and their parents. Child Development, 57, 20-33.

Bates, J.E., Olson, S.L., Pettit, G.S., & Bayles, K. (1982). Dimensions of individuality in the mother-infant relationship at six months of age. Child Development, 53, 446-461.

Belsky, J., Garduque, L, & Hrncir, E. (1984). Assessing performance, competence, and executive capacity in infant play: Relations to home environment and security of attachment. Developmental Psychology, 20, 406-417.

Berlin, L.J., Brooks-Gunn, J., Spiker, D., & Zaslow, M.J. (1995). Examining observational measures of emotional support and cognitive stimulation in black and white mothers of preschoolers. Journal of Family Issues, 16(5), 664-686.

Bradley, R.H. (1993). Children's home environments, health, behavior, and intervention efforts: A review using the HOME inventory as a marker measure. Genetic, Social, and General Psychology Monographs, 119(4), 439-490.

Bradley, R.H. (1981). The HOME inventory: A review of findings from the Little Rock Longitudinal Study. Infant Mental Health Journal, 198-205.

Bradley, R.H., & Caldwell, B.M. (1984a). The HOME inventory and family demographics. Developmental Psychology, 20(2), 315-320.

Bradley , R.H., & Caldwell, B.M. (1984b). 174 children: A study of the relationship between home environment and cognitive development during the first 5 years. In A.W. Gottfried (Ed.), Home environment and early cognitive development: Longitudinal research, (pp. 5-54) Orlando, FL: Academic Press, Inc.

Bradley, R.H. & Caldwell, B.M. (1984c). The relation of infants' home environments to achievement test performance in first grade: a follow up study. Child Development, 55, 803-809.

Bradley, R.H., & Caldwell, B.M. (1980). The relation of home environment, cognitive competence, and IQ among males and females. Child Development, 51, 1140-1148.

Bradley, R.H., & Caldwell, B.M. (1979). Home Observation for Measurement of the Environment: A revision of the preschool scale. American Journal of Mental Deficiency, 84(3), 235-244.

Bradley, R.H., Caldwell, B.M., Brisby, J., Magee, M., Whiteside, L., & Rock, S.L. (1992). The HOME inventory: A new scale for families of pre- and early adolescent children with disabilities. Research in Developmental Disabilities, 13, 313-333.

Bradley, R.H., Caldwell, B.M., & Rock, S.L. (1988). Home environment and school performance: A ten-year follow-up and examination of three models of environmental action. Child Development, 59, 852-867.

Bradley, R.H., Caldwell, B.M., Rock, S.L., Barnard, K.E., Gray, C., Hammond, M.A., Mitchell, S., Siegel, L., Ramey, C.T., Gottfried, A.W., &

Johnson, D.L. (1989). Home environment and cognitive development in the first three years of life: A collaborative study involving six sites and three ethnic groups in North America. Developmental Psychology, 25(2), 217-235.

Bradley, R.H., Mundfrom, D.J., Whiteside, L., Casey, P.H., & Barrett, K. (1994). A factor analytic study of the infant-toddler and early childhood versions of the HOME inventory administered to white, black, and Hispanic American parents of children born preterm. Child Development, 65, 880-888.

Bradley, R.H., Whiteside, L, Mundfrom, D.J., Casey, P.H., Kelleher, K.J., & Pope, S.K. (1994). Early indications of resilience and their relation to experiences in the home environments of low birthweight, premature children living in poverty. Child Development, 65, 346-360.

Caldwell, B.M., & Bradley, R.H. (1994). Environmental issues in developmental follow-up research. In S.L. Friedman & H.C. Haywood (Eds.), Developmental follow-up: Concepts, domains and methods (pp. 235-256), San Diego: Academic Press, Inc.

Caldwell, B.M., & Bradley, R.H. (1984). Home Observation for Measurement of the Environment. (Rev. ed.). (Administration Manual). Little Rock, Arkansas: University of Arkansas at Little Rock.

Chipuer, H.M., & Plomin, R. (1992). Using siblings to identify shared and non-shared HOME items. British Journal of Developmental Psychology, 10, 165-178.

Coll, C.T.G., Hoffman, J., & Oh, W. (1987). The social ecology and early parenting of Caucasian adolescent mothers. Child Development, 58, 955-963.

Desai, S., Michael, R.T., & Chase-Lansdale, P.L. (1990). The home environment: A mechanism through which maternal employment affects child development. Working Paper #20. New York: The Population Council.

Dunn, L.M. and Dunn, L.M. (1981). Peabody Picture Vocabulary Test - Revised: Manual. Circle Pines, MN: American Guidance Service.

Elardo, R.D., Bradley, R.H. & Caldwell, B.M. (1975). The relation of infants' home environments to mental test performance from six to thirty-six months: A longitudinal analysis. Child Development, 46, 71-76.

Garrett, P., Ng'andu, N., & Ferron, J. (1994). Poverty experiences of young children and the quality of their home environments. Child Development, 65, 331-345.

Gottfried, A.W. & Gottfried, A.E. (1984). Home environment and cognitive development in young children of middle-socioeconomic-status families. In A.W. Gottfried (Ed.), Home environment and early cognitive development: Longitudinal research, (pp. 57-115) Orlando, FL: Academic Press, Inc.

Hannan, K., & Luster, T. (1991). Influence of parent, child, and contextual factors on the quality of the home environment. Infant Mental Health Journal, 12(1), 17-30.

Hollenbeck, A.R. (1978). Early infant home environments: Validation of the Home Observation for Measurement of the Environment Inventory Developmental Psychology, 14(4), 416-418.

Johnson, D.L., Breckenridge, J.N. & McGowan, R.J. (1984). Home environment and early cognitive development in Mexican-American children. In A.W. Gottfried (Ed.), Home environment and early cognitive development: Longitudinal research, (pp. 151-195) Orlando, FL: Academic Press, Inc.

Klebanov, P.K., Brooks-Gunn, J., & Duncan, G.J. (1994). Does neighborhood and family poverty affect mothers' parenting, mental health, and social support? Journal of Marriage and the Family, 56, 441-455.

Kurtz, B.E., Borkowski, J.G., & Deshmukh, K. (1986). Metamemory and learning in Maharashtrian children: Influences from home and school. (Rev. ed.). Journal of Genetic Psychology, 149(3), 363-376.

Luster, T., & Dubow, E. (1992). Home environment and maternal intelligence as predictors of verbal intelligence: A comparison of preschool and school-age children. Merrill-Palmer Quarterly, 38(2), 151-175.

Matheny, A.P. Jr., Wilson, R.S., & Thoben, A.S. (1987). Home and mother: Relations with infant temperament. Developmental Psychology, 23(3), 323-331.

Menaghan, E.G., & Parcel, T.L. (1991). Determining children's home environments: The impact of maternal characteristics and current occupational and family conditions. Journal of Marriage and the Family, 53, 417-431.

Metzl, M.N. (1980). Teaching parents a strategy for enhancing infant development. Child Development, 51, 583-586.

Mitchell, S.K., & Gray, C.A. (1981). Developmental generalizability of the HOME inventory. Educational and Psychological Measurement, 41, 1001-1010.

Moore, K.A., & Snyder, N.O. (1991). Cognitive attainment among firstborn children of adolescent mothers. American Sociological Review, 56, 612-624.

Moore, K.A., Zaslow, M.J., Coiro, M.J., Miller, S.M., & Magenheim, E.B. (1995). How well are they faring? AFDC families with preschool-age children at the outset of the JOBS evaluation. Washington D.C.: U.S. Department of Helath and Human Services and U.S. Departent of Education.

Parcel, T.L., & Menaghan, E.G. (1989). Child home environment as a mediating construct between SES and child outcomes. Columbus, Ohio: The Ohio State University, Department of Sociology.

Parks, P.L., & Bradley, R.H. (1991). The interaction of home environment features and their relation to infant competence. Infant Mental Health Journal, 12(1), 3-16.

Parks, P.L., & Smeriglio, V.L. (1986). Relationships among parenting knowledge, quality of stimulation in the home and infant development. Family Relations, 35, 411-416.

Quint, J.C., Polit, D.F., Bos, H., & Cave, G. (1994). Interim findings on a comprehensive program for disadvantaged young mothers and their children. New York, NY: Manpower Demonstration Research Corporation.

Ramey, C.T., Yeates, K.O., & Short, E.J. (1984). The plasticity of intellectual development: Insights from preventive intervention. Child Development, 55, 1913-1925.

Siegel, L.S. (1984). Home environmental influences on cognitive development in preterm and full-term children during the first 5 years. In A.W. Gottfried (Ed.), Home environment and early cognitive development: Longitudinal research, (pp.197-233) Orlando, FL: Academic Press, Inc.

Stevenson, M.B., & Lamb, M.E. (1979). Effects on infant sociability and the caretaking environment on infant cognitive performance. Child Development, 50, 340-349.

Sugland, B.W., Zaslow, M., Smith, J.R., Brooks-Gunn, J., Coates, D., Blumenthal, C., Moore, K.A., Griffin, T., & Bradley, R.H. (1995). The early childhood HOME inventory and HOME-short form in differing racial/ethnic groups. Are there differences in underlying structure, internal consistency of subscales, and patterns of prediction? Journal of Family Issues, 16(5), 632-663.

Van Doominck, W.J., Caldwell, B.M., Wright, C., & Frankenburg, W.K. (1981). The relationship between twelve-month home stimulation and school achievement. Child Development, 52, 1080-1083.

Widmayer, S.M., Peterson, L.M., Larner, M., Carnahan, S., Calderon, A., Wingerd, J., & Marshall, R. (1990). Predictors of Haitian-American infant development at twelve months. Child Development, 61, 410-415.

Yeates, K.O., MacPhee, D., Campbell, F.A., & Ramey, C.T. (1983). Maternal IQ and home environment as determinants of early childhood intellectual competence: A developmental analysis. Developmental Psychology, 19(5), 731-739.