

## 2, 3세 유아의 사회 정서 문제 조기발견: ASQ:SE 선별 평가서의 표준화 연구

### Early Identification of 2- and 3-Year-Old Children for Social and Emotional Problems: A Preliminary Study of the Ages and Stages Questionnaires: Social-Emotional (ASQ:SE)

허 계 형\*  
Heo, kay Heung, Ph.D\*

#### ABSTRACT

본 연구 논문은 영 유아를 대상으로 한 사회 정서 문제의 선별 평가서인 Ages and Stages Questionnaire: Social-Emotional (ASQ:SE)의 표준화(신뢰도 및 타당도) 연구이다. 특히, ASQ:SE 질문지 가운데에서도 24개월과 36개월용을 이용하여 ASQ:SE의 내적 일치도(internal consistency reliability), 재검사 신뢰도(test-retest reliability), 절선 점수(cutoff points), 공인 타당도(concurrent validity)에 대해 연구하였다. 전체 447명의 부모가 참여한 가운데, 237명은 24개월용 질문지를 작성했고, 210명은 36개월용 질문지를 작성했다. 내적 일치도는 24개월 질문지에서 .71, 36개월 질문지에서는 .73이었다. 재검사 신뢰도는 24개월 질문지에서 100%, 36개월 질문지에서 97%이었다. 마지막으로 공인 타당도는 24개월과 36개월 질문지에서 95%이었다. 끝으로, 추후 다양한 대상으로 연구될 것이 추천되며 또한 이 연구에서 제외되었던 연령의 ASQ:SE 질문지의 내적 일치도, 신뢰도 및 타당도가 연구되어야 할 것이다.

---

\*우석대학교 유아특수교육과 전임강사

## I. INTRODUCTION

As the 21st century begins, the world faces many factors that pose significant risk to normal family functioning (Hunt, Johnson, Owne, Ormerod, & Babbitt, 1990; Walker, Colvin, & Ramsey, 1995). Of particular concern is the increase in violence against children, and resulting violent behavior of youth. Many studies of young violent offenders conclude that the origins of socially and emotionally problematic behavior begin in early childhood (Cicchetti, 1993; Hunt et al., 1990; Patterson, Reid, & Dishion, 1992). It is generally agreed that social and emotional problems that begin early in a child's life tend to increase in intensity as a child grows older and these problems often become more serious over time (Walker et al., 1995; Yoshikawa, 1994). Research by Feil, Walker, and Severson (1995) suggests that society's efforts at remediation would be most effective concentrated on prevention aimed at families and young children. They and many other researchers who study antisocial youth support prevention and early intervention, reasoning that the further a youth moves along the progression of social and emotional difficulties, the greater the risk for later behavior problems. However, even though the importance of early identification has been emphasized in recent literature, many young children who are at-risk for developing social and emotional problems are still not being identified in a timely fashion (Bennett, Nickel, Squires, & Woodward, 1997; Forness,

Kavale, MacMillan, Asanow, & Duncan, 1996; Little, Murphy, Bishop, & Arnett, 1994).

One likely explanation of the failure to identify social and emotional problems early is the lack of clarity surrounding the definitions of social and emotional competence in very young children (Mash & Dozois, 1996). Wittmer, Doll, and Strain (1996) address the difficulties of measuring the social and emotional area in very young children because of the lack of comprehensive definitions of social and emotional development. These researchers have chosen to address the definition problem by identifying social and emotional development as separate but reciprocal areas of competence and disability. They define social competence as the externalizing behaviors, attitudes and affect that a child brings to social interactions, including the areas of cooperative and pro-social behavior, active initiation and maintenance of peer and adult relationships, and managing aggression and conflict. Wittmer and her colleagues define emotional competence as the internal emotions and their overt manifestations including self-worth, sense of mastery, and emotional regulation and reactivity.

Another major barrier to identification of social and emotional problems in young children is the lack of psychometrically sound, low-cost assessments that can accurately identify the lack of social and emotional competency of infants, toddlers, and preschool-age children (Wittmer et al., 1996; Yockelson, 1998). Researchers (e.g., Sexton, Thompson, Perez, & Rheams, 1990) have pointed out that the use of appropriate

standardized instruments is critical because incorrect conclusions and decisions may be made when measures that are poorly constructed and studied are used. Studies of many existing instruments for the social and emotional area have found that most scales are severely limited in their technical adequacy, especially regarding validity and reliability (Bracken, Keith, & Walker, 1994). Social and emotional tools that are currently used also have limited data on norms and determination of cutoffs (Little, Murphy, Bishop, & Arnett, 1994).

Another barrier to the early identification of young children with social and emotional problems is the size of the population that needs to be assessed. Many researchers agree that there are significant numbers of children in need of intervention for social and emotional problems (Mash & Dozois, 1996). Conducting frequent screenings on the large numbers of children who are suspected of having social and emotional problems can make the cost of screening high, especially if screening is conducted by highly trained professionals. Effective screening tools need to be quick, simple, and affordable in order to differentiate large numbers of typically developing young children from those who are at risk or who have problems (McLean, Bailey, & Wolery, 1996). Additionally, the prevalence rate of social and emotional problems in young children varies with the nature of the disorder, the child's age, social class, and the criteria used to define the problem. Screening of a risk population for social and emotional development is recommend to be

conducted in intervals over time because of the nature of development and of the inception of problems. An example of a very successful screening system in the United States is the Ages and Stages Questionnaires (ASQ) (Bricker & Squires, 1999), a parent-completed developmental questionnaire. Many U.S. cities conduct large child-find programs for children with developmental delays with the ASQ. Studies of the ASQ have found that the cost of screening a child three times a year including professional time, protocols, and agency overhead, range from \$7.50 to \$25 yearly, depending on the agency (Chan, 1997).

This article intends to add to the body of knowledge about a promising screening tool designed to fill the need for efficient, valid, reliable, and low-cost early identification of social and emotional problems in young children. This study examine the preliminary reliability and validity of the Ages and Stages Questionnaires: Social-Emotional (ASQ:SE), a recently developed parent-completed screening tool to identify social and emotional problems in very young children. The ASQ:SE was developed by the Early Intervention Program, Center on Human Development, University of Oregon, and has been specifically designed to screen large number of children using a parent-completed, family friendly, and cost-efficient format.

Normative data for the ASQ:SE are currently being collected; the research described here investigated selected psychometric properties of the ASQ:SE for questionnaires at the 2-year-old (24-month) and 3-year-old (36-month) intervals.

The study was designed to investigate selected psychometric properties of the ASQ:SE, including validity and reliability.

## II. METHOD

### 1. Subjects

Four hundred forty-seven parents participated in the study along with 447 children. Two hundred thirty-seven children were between the ages of 21 and 27 months, and 210 were between the ages of 33 and 42 months. Among these, fifty-eight parents from the 24-month age group and sixty-three parents from the 36-month age group also voluntarily participated in the validity and reliability studies.

Parents were contacted and recruited by a variety of methods in order to assure an economically diverse group of participants. Parents were recruited through birth announcements in newspapers published in major cities in Oregon; USA through day care providers and preschools; through public places where children and parents gather, such as shopping malls and playgrounds; and through agencies that provide support and information on parenting young children. All participation was voluntary and all information provided by participants was kept confidential. An effort was made to recruit a diverse sample of parents and young children that was representative of race, gender, and educational level in distributions similar to that of the U.S. population

(US Bureau of the Census, 1998).

### 2. Tests and Measures

Two measures were used: the Ages and Stages Questionnaires: Social-Emotional (ASQ:SE), and the Child Behavior Checklist for Ages 2-3 (CBCL-2/3) (Achenbach, 1992). In addition, a demographic form and a research opportunity form were used in the study.

The Ages and Stages Questionnaires: Social-Emotional (ASQ:SE). The ASQ:SE was developed after an extensive review of the literature of existing diagnostic assessments to select behaviors representative of the social and emotional domains. Items were reviewed by experts in diverse disciplines such as special education, early childhood education, school psychology, and communication disorders and sciences, and by representatives from mental health, social welfare and child protection agencies. To date, information has been gathered and summarized on the developmental appropriateness, age/content appropriateness, family friendliness, redundancy, and clarity of items for all age intervals.

Although it was an original goal to word all items in a positive way (e.g., "Is your baby happy?"), the final version of the ASQ:SE includes both positively- and negatively-stated questions. Even though positively-stated questions were more family-friendly and acceptable to parents, they were not always appropriate to target behaviors. Therefore, negatively-stated questions were included (e.g., "Does your child have eating

problems such as stuffing foods, vomiting, or eating non-food items?"). The ASQ:SE consist of eight intervals for assessing children ages 3 months to 60months, with questionnaires at 6, 12, 18, 24, 30, 36, 48, and 60 months. The length of the questionnaires vary, depending on the age of the child (e.g., 24 months -- 28 items; 36 months -- 34 items). Each questionnaire consists of questions about the child's relationship with caregivers; self-regulation; communication; coping/autonomy; relationship with peers (on 24- through 60-month questionnaires); and general questions about family concerns and worries.

Each questionnaire can be completed by a parent or a caregiver (e.g., foster parent) in 10-20 minutes, depending on the length of the questionnaire and the time it takes for parents to read and mark the appropriate answers. Reading level is approximately fifth to sixth grade. For parents who do not read at a fifth grade level, the questionnaires can be used as an interview tool.

Parents respond to each question by checking one of three choices: "most of the time," "sometimes," or "never or rarely." Additionally, parents identify their level of concern for each item by checking "yes" or "no" under a column heading "Is this a concern?" Scoring is completed by summing the numeric equivalencies for each response, and the number of yes responses from the concern column. Points are awarded to questions as follows: 0 points if a negative behavior never or rarely occurs or if a positive behavior occurs most of time; 5 points if a behavior occurs sometimes; and 10 points if a negative behavior occurs most of time

or if a positive behavior never or rarely occurs. Five points are added for each yes response from the parental concern column.

The Child Behavior Checklist for Ages 2-3 (CBCL/ 2-3) (Achenbach, 1992). CBCL/ 2-3 was used as a concurrent measure in this study because it is regarded as one of the best current assessments of behavior problems in children (McConaughy, 1992). The version of the CBCL for 2-3-year-olds was revised in 1992 and contains 99 questions plus one open-ended item. Scoring options on the CBCL/ 2-3 are "not true," "somewhat or sometimes true," and "very true or often true," and are the options are scored respectively 0, 1, 2. Separate scores can be obtained for externalizing or internalizing behaviors or for more specific disorders within those categories such as anxious, depressed, aggressive, or destructive. Additionally, summary scores representing "internalizing" or "externalizing" problems can be calculated (Achenbach, 1992). Cutoff points were obtained for this instrument by taking demographically-matched referred and non-referred samples and determining the best cut-off points between these groups that would minimize the number of over-referrals (Achenbach, 1992)

Demographic Form. The demographic form for this study asked for general family information on the child's gender, date of birth, mother's age at child's birth, child's birth weight, family income, caregiver's level of education, racial group, and disability status. This information was collected in order to recruit a stratified sample.

Research Opportunity. Subjects were invited to

participate in the reliability and validity studies after they had completed the ASQ:SE. When they returned their questionnaire, they were mailed an information sheet titled "Research Opportunities" and were asked if they wanted to participate further in the validity and reliability study. This information was used to ensure voluntary participation for the validity and reliability study.

### 3. Procedures

Parents were recruited in a variety of ways, including birth announcements in newspapers, booths at shopping malls, and local preschools and daycare centers, and professional contacts with agencies serving children with diagnosed disabilities. Once parents indicated a willingness to participate, they were given a packets of materials containing a consent form, an ASQ:SE questionnaire, a demographic form, and research opportunity form. This packet was distributed by research personnel directly to parents (e.g., at shopping malls), by mail (e.g., parents contacted through birth announcements), and through preschool teachers. Packets were also mailed to personnel in agencies serving young children and families who then distributed them to interested parents. In most cases, parents returned their completed packets to the research staff by mail in a pre-stamped envelope. Once the packets were received, parents who returned the research opportunity form were contacted. These parents were then mailed the CBCL/ 2-3 (for the validity study) and a second ASQ:SE (for the test-retest reliability study) with

instructions to complete and return the forms in a pre-stamped enveloped within 1 week. All data were coded and labeled with identification numbers, and standard practices for maintaining confidentiality were used.

## RESULT

### 1. Population

The 24- and 36-month intervals of the ASQ:SE were the object of this study. Children between the ages of 21 and 27 months were used for the study of the 24-month interval, and children between the ages of 33 and 45 months were used for the 42-month interval of the ASQ:SE. Approximately 9% of the participants were recruited through birth announcement in newspapers; 11% through newspaper advertisements; 25% through agency personnel who attended national conferences and agreed to field test the ASQ:SE; 16% at shopping malls; and 39% through early intervention centers and parent education programs. Demographic information on participants is presented in Table 1. Ethnicity of children is presented in Table 2. Since parents did not answer all demographic questions, data are missing in some categories.

&lt;Table 1&gt; Demographic Characteristics of Participants

	24-month		36-month	
	N	%	N	%
<b>Gender</b>				
Male	121	51.1	96	46
Female	114	48.1	114	54
Missing	2	0.8	0	0
Total	237	100	210	100
<b>Education Level of Mother</b>				
Less than high school	29	12.2	12	5.7
High school	60	25.3	51	24
AA degree	73	31	62	30
4 year college of above	70	30	70	33
Don't know	1	0.4	1	0.48
Missing	4	1.7	14	6.7
Total	237	100	210	100
<b>Family Income</b>				
0-\$12,000	62	26	40	19
\$12,000-24,000	36	15	40	19
\$24,000-40,000	51	22	42	20
Over 40,000	65	27	71	34
No response	18	8	5	2.4
Missing	5	2	12	5.7
Total	237	100	210	100
<b>Type of Respondent</b>				
Mother	218	92.4	173	82.3
Father	8	3.4	16	7.6
Grandparents	6	2.5	16	7.6
Guardian	3	1.3	2	1.0
Others	1	0.4	2	1.0
Missing	1	0.4	1	0.5
Total	237	100	210	100

## 2. Validity

The validity of a screening measure is determined by its ability to accurately discriminate between individuals in a given population who appear OK or normal on the variable of interest (i.e., not identified with a possible problem) and those individuals who appear as at risk, or abnormal (i.e., identified with a possible problem). In this study, the variable of interest was social-emotional competence. To determine the validity of the ASQ:SE, two steps were necessary. The purpose of the first step was to collect adequate normative data to permit the establishment of valid and reliable cutoff points. The purpose of second step was to examine the agreement between children screened by the ASQ:SE using the first step and the selected

concurrent measures -- in this case, a comparison of how children scored on the CBCL/ 2-3 and the ASQ:SE.

Step One: Establishing Cutoff Points. The cutoff point for a screening tool is the range of scores that determines whether a child should be recommended for further, more in-depth assessment. For the ASQ:SE, children whose total score is equal to or exceeds the established cutoff point should be referred for further evaluation of social and emotional competency (i.e., risk status with possible problem).

Children whose total score is below the cutoff point are not recommended for immediate evaluation, though this can be overridden by parent/ professional concerns. A child whose total score is near the cutoff point should be followed

<Table 2> Ethnicity of Children in the ASQ:SE Norming Sample (N = 447)

Ethnicity of Child	ASQ:SE Interval			
	24-Month Questionnaire		36-Month Questionnaire	
	n	Percent	n	Percent
African American	7	3.0	2	1
Asian	11	4.7	5	2.4
Caucasian	130	54.9	149	71
Hawaiian	15	6.3	4	1.9
Hispanic	3	1.3	7	3.3
Native American	1	0.4	2	1
Pacific Islander	3	1.3	2	1
Multiracial	48	20.2	13	6.2
Other	18	7.6	14	6.7
Missing	1	0.4	12	5.7
Total	237	100	210	100



more closely and receive further screening at regular intervals (Squires et al, 1996).

Sensitivity (i.e., the ability of the test to detect true positive or those children with developmental problems) was calculated for the ASQ:SE, as well as specificity (i.e., the ability of the test to exclude from further evaluation those children who are developing typically). A contingency table illustrating the possible relationships of the ASQ:SE with the criterion measure can be found in Table 3.

In order to determine the optimal cutoff point for each of the two intervals studied, five steps were employed. First, means, standard deviations, medians, and semi-interquartile ranges for the questionnaires under study were calculated. Second, cutoff points using standard deviations and means were examined at 1, 1-1/2, and 2 standard

deviations above the mean. Third, the cutoff points using the semi-interquartile ranges were examined at 1-1/2, 2-1/4, and 3 semi-interquartile ranges. Fourth, classification statistics (i.e., sensitivity, specificity, over-referral, and under-referral) were calculated using each of the potential cutoff scores in comparison to a concurrent measure, the CBCL/2-3. The results of the CBCL/2-3 were used to calculate optimal classification statistics (e.g., sensitivity, specificity, true positives) using the CBCL 2/3 cutoff score recommended in the manual (Achenbach, 1992). In the fifth step, the optimal cutoff points determined in step four for each interval were examined using generating Receiver Operating Characteristic (ROC) curves (Swets & Pickett, 1982) based on true positive and false positive scores. The optimal cutoff score, maximizing true positives and minimizing false

<Table 3> Possible Classifications on the ASQ:SE in Comparison to the CBCL/2-3

		CBCL/2-3	
		Risk (positive)	OK (negative)
ASQ:SE	Risk (positive)	a Correct Decision (true positive)	b Error (over-referral)
	OK (negative)	c Error (under-referral)	d Correct Decision (true negative)

Sensitivity (true positive) =  $a / a + c$

Specificity =  $d / b + d$

Over-referral =  $b / a + b + c + d$

Under-referral =  $c / a + b + c + d$

Percent agreement =  $a + d / a + b + c + d$

positives, was chosen for the each studied interval. Optimal cutoff scores were determined to be 77.5 (3 semi-interquartile range) for the 24-month questionnaire, and 85.6 (2-1/4 semi-interquartile range) for the 36-month questionnaire. A more detailed explanation for obtaining the cutoff scores, as well as the data obtained for the five steps described above can be found in the appendix.

Step Two: Establishing Concurrent Validity For developmental screening instruments, the classification of the screening tool is compared with a criterion measure or battery of measures. Based on the

agreement between the screening instrument and the criterion measures, the accuracy of the screening test is determined.

In order to investigate the accuracy of parents in completing questionnaires about their young children's social and emotional competencies, the reliability and validity of the ASQ:SE were examined. Concurrent validity was measured by comparing the classification of the children's performances on the ASQ:SE with classifications of children obtained in one of two ways. First, parents or caregivers completed a concurrent measure, the CBCL. Cutoff scores on the CBCL

<Table 4> Concurrent Validity: Classification Agreement Between the ASQ:SE and CBCL/2-3 at the 24-Month Interval (N = 58)

ASQ:SE		Risk	OK	Total	
	Risk	0	1		1
	OK	2	55		57
		2	56	58	

Sensitivity	Specificity	True Positive	False Positive	Over-Referral	Under-Referral
-	.98	-	.02	.2	.03

<Table 5> Concurrent Validity: Classification Agreement Between the ASQ:SE and CBCL/2-3 at the 36-Month Interval (N = 63)

ASQ:SE		Risk	OK	Total	
	Risk	4	2		6
	OK	1	56		57
		5	58	63	

Sensitivity	Specificity	True Positive	False Positive	Over-Referral	Under-Referral
.80	.97	.80	.03	.03	.16

were 64; on or above 64 indicated a child was classified as risk; below 64 as OK. Using the ASQ:SE cutoffs described above, a child's score was risk if it was on or above the cutoff score for each interval; OK if it was below the cutoff score. In the comparison of scores between instruments, one of four outcomes was possible: 1) Both test classified the child as OK; 2) both tests classified the child as at risk or in need of further testing; 3) the CBC: indicated the child as OK and the ASQ:SE indicated the child as at risk; and 4) the ASQ:SE classified the child as OK and the CBCL as at risk. Contingency tables showing agreement between the concurrent measure and the ASQ:SE by age interval are contained in Tables 4 and 5. The percent agreement for the both 24-month and 36-month questionnaires was 95%.

### 3. Reliability

Internal Consistency. Internal consistency was examined using coefficient alpha, or Cronbach's alpha (Cronbach, 1951). Coefficient alpha was calculated from the variances of individual test items and the variance of the total test scores for the ASQ:SE at each test interval. A Cronbach's alpha of .70 or higher is considered to reflect acceptable internal consistency (Nunnally, 1976). The Cronbach's coefficient alpha for the 24-month questionnaire was .71 ( $n = 134$ ), and for the 36-month questionnaire was .73 ( $n = 181$ ).

Test-Retest Reliability. A sub-sample of parents completed a second ASQ:SE within a one-to-three-week time period after returning the first

completed questionnaire to research personnel. The parents were blind to the results of the first ASQ:SE that they had completed. Test-retest reliability was determined by comparing the results of two ASQ:SE completed by parents.

Test-retest reliability was measured as percentage agreement between classifications (i.e., risk, OK) of the child based upon the questionnaires completed by parents at time 1 and time 2. (The procedures for determining cutoff scores for risk/ OK status is described in the Validity section above.) The agreement between the first and second completions of the 24-month ASQ:SE was 100% and on the 36-month questionnaire was 97%.

## IV. DISCUSSION

The primary goal for this study was to investigate selected psychometric properties of two intervals of the Ages and Stages Questionnaires: Social-Emotional (ASQ:SE). Reliability and validity findings for the 24-month and 36-month intervals of the ASQ:SE are promising. Internal consistency was .71 for the 24-month and .73 for the 36-month questionnaires, both within acceptable levels. Test-retest reliability agreement of 100% for the 24-month ASQ:SE and 97% for the 36-month ASQ:SE was high. The Pearson product-moment correlation was also high;  $r$  was .89 for the 24-month interval and .92 for the 36-month interval. Concurrent validity agreement of 95% for both the 24-month and 36-month questionnaires was also promising.

## 1. Contributions of the Study

The current study of the 24- and 36-month intervals of the ASQ:SE contributes to the research on identification of social problems in young children in the following ways. First, the findings of this study add support to the psychometric base (i.e., reliability and validity) of this parent-completed social and emotional screening instrument. Second, the study findings support the growing body of literature that attempts to define social and emotional competence in young children.

This study contributes to fulfilling the urgent need for a standardized and population-specific screening instrument for young children with social and emotional problems. The current study includes a minimum of 150 cases at each interval studied, with an attempt made to collect a representative group of 2- and 3-year-olds. The study has shown that the under-referral and over-referral rates for the 24- and 36-month ASQ:SE are acceptable, indicating that in most cases the ASQ:SE will accurately identify children who are in need of further evaluation and those who are not.

The current study of the ASQ:SE contributes to the growing literature that attempts to identify and define social and emotional development. The ASQ:SE has been developed with consideration to identifying social and emotional competence with variables such as time and settings; individual development; and individual, family and cultural values. The ASQ:SE has a format that is specifically designed for very young children, with intervals

from 6 to 60 months. There are few social and emotional instruments that are specifically developed for young children and that span the entire preschool years.

The current study supports the utility of the ASQ:SE as an easy-to-use and affordable parent-completed screening test. The ASQ:SE is relatively inexpensive to administer because it relies on parents or other primary caregivers to complete simple, easy-to-read questionnaires on their children. Parents expressed informally throughout the study that they feel the ASQ:SE is easy to use and easy to understand. The format of the ASQ:SE that includes and considers parental concerns in the scoring system is innovative. Studies have found that parents' concerns about their children's behavior can predict future problems (Mulhern et al., 1994). However, few instruments specifically request information about parental concerns.

In sum, the study of the 24- and 36-month ASQ:SE contributes to alleviating the significant need for a psychometrically sound, low-cost screening tool that can accurately reflect the emotional and social competencies of young children. The ASQ:SE system can assist parents and caregivers, and early intervention and early childhood personnel in the timely identification of children whose behaviors indicate future difficulties in the social and emotional areas.

## 2. Limitations of the Study

The primary limitation of this study pertains to the reliance upon a single type of informant for a

portion of the concurrent validity studies. That is, parents or caregivers were used as the primary screeners of children's social and emotional competence, both for the CBCL, the concurrent validity measure, and for the ASQ:SE. In addition, the CBCL is not a "gold standard" for measuring problem behaviors (Doll, 1998) and will not always give an accurate assessment of the presence of behavior problems. Additional methods of gathering information such as clinical observations and professionally administered assessments will increase the external validity of this study.

Second, only selected intervals of the ASQ:SE were examined in this study. While findings provided promising evidence in support of the psychometrics of the intervals studied, they must be viewed as only a partial study of the ASQ:SE screening system. Studies have found that children's behavior changes with different settings and social partners, and also changes over time (Achenbach et al., 1987). It will be beneficial to investigate the validity and reliability of the ASQ:SE over time as well as complete psychometric investigations of the rest of the intervals. Earlier intervals may prove more challenging to study, especially during the first year of life (i.e., 6-, 12-, 18-month intervals), because of transient problems in state and arousal experienced by some infants.

Third, characteristics of the sample may also have presented an additional limitation. Participating children were mostly typically developing and may have been easier to assess. Greater ethnic diversity and additional children with social and emotional disabilities will increase validity of these findings.

Further studies are also needed to examine whether subgroups of parents, especially those in high risk environments, can be accurate assessors of their young children's behavior.

Although the importance of early identification has been emphasized, many young children with disabilities, or who are at-risk for developing disabilities, are still not being identified in a timely manner. The lack of psychometrically sound, low-cost screening tools that can reflect the social and emotional competencies of young children is one of the major barriers to the current lack of early identification. An accurate screening tool can help to improve the quality of life for young children by identifying them in a timely manner for further assessment and early intervention services. The development of the ASQ:SE is one such strategy that appears to have great potential.

## REFERENCES

- Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, 101, 213-232.
- Achenbach, T. M. (1992). *Child behavior checklist for ages 2-3 (CBCL/2-3)*. Burlington, VT: University of Vermont, Department of Psychiatry.
- Bennett, F., Nickel, R., Squires, J., & Woodward, B. (1997). Developmental screening/surveillance. In H. Wallace, R. Biehl, R. MacQueen, & J. Blackman (Eds.), *Children with disabilities and*

- chronic illness* (pp. 236-247). Oakland, CA: Mosby Yearbook.
- Bracken, B., Keith, L., & Walker, K. (1994). Assessment of preschool behavior and social-emotional functioning: A review of thirteen third-party instruments. *Assessment in Rehabilitation and Exceptionality*, 1(4), 331-346.
- Bricker, D., & Squires, J., with Mounts, L., Potter, L., Nickel, R., Twombly, E., & Farrell, J. (1999). *Ages and stages questionnaires: A parent-completed, child-monitoring system* (2nd ed). Baltimore: Paul Brookes.
- Chan, B. (1997). *Ages & Stages cost analysis*. Minneapolis, MN: Minnesota Department of Education.
- Cicchetti, D. J. (1993). Developmental psychopathology: Reactions, reflections, projections. *Developmental Review*, 13, 471-502.
- Doll, B. (1998). Review of the Child Behavior Checklists. In J. C. Impara & B. S. Plake (Eds.), *The thirteenth mental measurements yearbook* (pp. 217-220). Lincoln, NE: University of Nebraska Press.
- Feil, E. G., Walker, H. M., & Severson, H. H. (1995). The early screening project for young children with behavior problems. *Journal of Emotional and Behavioral Disorders*, 3(4), 194-202.
- Forness, S., Kavale, K., MacMillan, D., Asanow, J., & Duncan, B. (1996). Early detection and prevention of emotional or *behavioral disorders*: Developmental aspects of systems of care. *Behavioral Disorders*, 21(3), 226-240.
- Hunt, F. M., Johnson, C. R., Owen, G., Ormerod, A. J., & Babbitt, R. L. (1990). Early intervention for severe behavior problems: The use of judgment based assessment procedures. *Topics in Early Childhood Special Education*, 10(3), 111-121.
- Little, M., Murphy, J. M., Bishop, S. J., & Arnett, H. L. (1994). Screening 4 and 5 old children for psychosocial dysfunction: A preliminary study with the Pediatric Symptom Checklist. *Journal of Developmental and Behavioral Pediatrics*, 15(3), 191-197.
- Mash, E. J. & Dozois, D. J. (1996). Child psychopathology: A developmental systems perspective. In E. J. Mash & R. A. Barkley (Eds.), *Child psychopathology* (pp. 3-60). New York: Guilford.
- McLean, M., Bailey, D. B., & Wolery, M. (1996). *Assessing infants and preschoolers with special needs*. Columbus, OH: Merrill.
- Squires, J., Nickel, R., Eisert, D. (1996). Early detection of developmental problems: Strategies for monitoring young children in practical settings. *Developmental and Behavioral Pediatrics*, 17(6), 420-427.
- Mulhern, S., Dworkin, P. H., & Bernstein, B. (1994). Do parental concerns predict a diagnosis of attention-deficit hyperactivity disorder? *Journal of Developmental and Behavioral Pediatrics*, 15(5), 348-352.
- Patterson, G. R., Reid, J. B., & Dishion, T. J. (1992). *Antisocial boys*. Eugene, OR: Castalia.
- Sexton, D., Thompson, B., Perez, J., & Rheams, T. (1990). Maternal versus professional estimates of developmental status for young children with handicaps: An ecological approach. *Topics in Early Childhood Special Education*, 10(3), 80-95.
- U.S. Bureau of the Census. (1998). *United States population estimates, by age, race, and Hispanic origin, 1990-1997* (Release PPL-91). Washington, DC: U.S. Bureau of the Census, Population Division.
- Walker, H. M., Colvin, G., & Ramsey, E. (1995). *Antisocial behavior in school: Strategies and best*

- practice*. Practice Grove: Brooks/Cole.
- Wittmer, D., Doll, B., & Strain, P. (1996). Social and emotional development in early childhood: The identification of competence and disabilities. *Journal of Early Intervention*, 20(4), 299-318.
- Yockelson, S. J. (1998). Identification of infants and young children with social-emotional problems. *Infants and Young Children*, 10(3), 1-9
- Yoshikawa, H. (1994). Prevention as cumulative protection: Effects of early family support and education on chronic delinquency and its risks. *Psychological Bulletin*, 115(1), 28-54.

## Appendix: Determining Cutoff Points

In order to determine the optimal cutoff points for each of the two intervals studied, five steps were employed.

Step 1. Means, standard deviations, medians, and semi-interquartile ranges for each questionnaire were determined. The results are shown in Table 1.

Step 2. The cutoff points determined using standard deviations and means were examined. The specific cutoff points studied were 1, 1½, and 2 standard deviations above the mean.

Step 3. The cutoff points using semi-interquartile ranges were examined. Results of this analysis comparing cutoff points using standard deviations and semi-interquartile ranges are shown in Table 2.

Step 4. Classification statistics (i.e., sensitivity, specificity, over-referral, and under-referral) were calculated using each of the potential cutoff scores. In order to determine classification statistics, a concurrent measure (i.e., the Child Behavior

<TABLE 1> Means, Standard Deviations, Medians, and Interquartile Ranges of the 24- and 36-Month ASQ:SE

ASQ:SE Interval	n	Mean	Standard Deviation	Median	Interquartile Range
24-Month	237	32.23	29.38	25	35
36-Month	214	46.82	42.17	35	45

<TABLE 2> ASQ:SE Cutoff Points Derived Using Standard Deviations and Semi-Interquartile Ranges

ASQ:SE Interval	Standard Deviation			Semi-Interquartile Range		
	1	1½	2	1½	2¼	3
24-Month	62.62	76.3	91	51.25	64.38	77.5
36-Month	88.99	110.08	131.16	68.75	85.63	102.5

<TABLE 3> Contingency Tables Comparing Classifications on the ASQ:SE and CBCL/2-3 for the 24-Month Questionnaire

	Standard Deviation(SD)				Semi-Interquartile Range(S-IQR)				
I SD	CBCL/2-3				1 ½ S-IOR	CBCL/2-3			
	ASQ:SE	Risk	OK	Total		ASQ:SE	Risk	OK	Total
		0	3	3			1	6	7
	2	53	55		1	50	51		
	Total	2	56	58	Total	2	56	58	
1 ½ SD	CBCL/2-3				2 ¼ S-IOR	CBCL/2-3			
	ASQ:SE	Risk	OK	Total		ASQ:SE	Risk	OK	Total
		0	1	1			0	3	3
	2	55	57		2	53	55		
	Total	2	56	58	Total	2	56	58	
2 SD	CBCL/2-3				3 S-IOR	CBCL/2-3			
	ASQ:SE	Risk	OK	Total		ASQ:SE	Risk	OK	Total
		0	1	1			0	1	1
	2	55	57		2	55	57		
	Total	2	56	58	Total	2	56	58	

<TABLE 4> Classification Statistics by Standard Deviation Units(SD) and Semi-Interquartile Ranges(S-IQR) for the 24-Month Questionnaire (N=58)

		Sensitivity	Specificity	True Positive	False Positive	Over-Referral	Under-Referral
SD	1	-	.95	-	.05	.05	.03
	1 ½	-	.98	-	.02	.02	.03
	2	-	.98	-	.02	.02	.03
S-IQR	1 ½	.50	.89	.50	.11	.10	.02
	2 ¼	-	.95	-	.05	.05	.03
	3	-	.98	-	.02	.02	.03



<TABLE 5> Contingency Tables Comparing Classifications on the ASQ:SE and CBCL/2-3 for the 36-Month Questionnaire

	Standard Deviation(SD)	Semi-Interquartile Range(S-IQR)																																				
1 SD	<p>CBCL/2-3</p> <table border="1"> <tr><td></td><td>Risk</td><td>OK</td><td>Total</td></tr> <tr><td>ASQ:SE</td><td>Risk</td><td>4</td><td>2</td><td>6</td></tr> <tr><td></td><td>OK</td><td>1</td><td>56</td><td>57</td></tr> <tr><td>Total</td><td>5</td><td>58</td><td>63</td></tr> </table>		Risk	OK	Total	ASQ:SE	Risk	4	2	6		OK	1	56	57	Total	5	58	63	<p>1 ½ S-IOR</p> <p>CBCL/2-3</p> <table border="1"> <tr><td></td><td>Risk</td><td>OK</td><td>Total</td></tr> <tr><td>ASQ:SE</td><td>Risk</td><td>4</td><td>7</td><td>11</td></tr> <tr><td></td><td>OK</td><td>1</td><td>51</td><td>52</td></tr> <tr><td>Total</td><td>5</td><td>58</td><td>63</td></tr> </table>		Risk	OK	Total	ASQ:SE	Risk	4	7	11		OK	1	51	52	Total	5	58	63
	Risk	OK	Total																																			
ASQ:SE	Risk	4	2	6																																		
	OK	1	56	57																																		
Total	5	58	63																																			
	Risk	OK	Total																																			
ASQ:SE	Risk	4	7	11																																		
	OK	1	51	52																																		
Total	5	58	63																																			
1 ½ SD	<p>CBCL/2-3</p> <table border="1"> <tr><td></td><td>Risk</td><td>OK</td><td>Total</td></tr> <tr><td>ASQ:SE</td><td>Risk</td><td>2</td><td>0</td><td>2</td></tr> <tr><td></td><td>OK</td><td>3</td><td>58</td><td>61</td></tr> <tr><td>Total</td><td>5</td><td>58</td><td>63</td></tr> </table>		Risk	OK	Total	ASQ:SE	Risk	2	0	2		OK	3	58	61	Total	5	58	63	<p>2 ¼ S-IOR</p> <p>CBCL/2-3</p> <table border="1"> <tr><td></td><td>Risk</td><td>OK</td><td>Total</td></tr> <tr><td>ASQ:SE</td><td>Risk</td><td>4</td><td>2</td><td>6</td></tr> <tr><td></td><td>OK</td><td>1</td><td>56</td><td>57</td></tr> <tr><td>Total</td><td>5</td><td>58</td><td>63</td></tr> </table>		Risk	OK	Total	ASQ:SE	Risk	4	2	6		OK	1	56	57	Total	5	58	63
	Risk	OK	Total																																			
ASQ:SE	Risk	2	0	2																																		
	OK	3	58	61																																		
Total	5	58	63																																			
	Risk	OK	Total																																			
ASQ:SE	Risk	4	2	6																																		
	OK	1	56	57																																		
Total	5	58	63																																			
2 SD	<p>CBCL/2-3</p> <table border="1"> <tr><td></td><td>Risk</td><td>OK</td><td>Total</td></tr> <tr><td>ASQ:SE</td><td>Risk</td><td>2</td><td>0</td><td>2</td></tr> <tr><td></td><td>OK</td><td>3</td><td>58</td><td>61</td></tr> <tr><td>Total</td><td>5</td><td>58</td><td>63</td></tr> </table>		Risk	OK	Total	ASQ:SE	Risk	2	0	2		OK	3	58	61	Total	5	58	63	<p>3 S-IOR</p> <p>CBCL/2-3</p> <table border="1"> <tr><td></td><td>Risk</td><td>OK</td><td>Total</td></tr> <tr><td>ASQ:SE</td><td>Risk</td><td>2</td><td>0</td><td>2</td></tr> <tr><td></td><td>OK</td><td>3</td><td>58</td><td>61</td></tr> <tr><td>Total</td><td>5</td><td>58</td><td>63</td></tr> </table>		Risk	OK	Total	ASQ:SE	Risk	2	0	2		OK	3	58	61	Total	5	58	63
	Risk	OK	Total																																			
ASQ:SE	Risk	2	0	2																																		
	OK	3	58	61																																		
Total	5	58	63																																			
	Risk	OK	Total																																			
ASQ:SE	Risk	2	0	2																																		
	OK	3	58	61																																		
Total	5	58	63																																			

<TABLE 6> Classification Statistics by Standard Deviation Units(SD) and Semi-Interquartile Ranges(S-IQR) for the 36-Month Questionnaire (N=63)

		Sensitivity	Specificity	True Positive	False Positive	Over-Referral	Under-Referral
SD	1	.80	.97	.80	.03	.03	.16
	1 ½	.40	1.0	.40	0	0	.48
	2	.40	1.0	.40	0	0	.48
S-IQR	1 ½	.80	.88	.80	.12	.11	.16
	2 ¼	.80	.97	.80	.03	.03	.16
	3	.40	1.0	.40	0	0	.48

Checklist for Ages 2-3 [CBCL 2/3] [Achenbach, 1992]) was used. That is, classifications of the child's performance on the ASQ:SE and CBCL/2-3 were compared across cutoff scores. A cutoff score of 64 on the CBCL/2-3 was used for both ASQ:SE intervals. This score was chosen because the score was recommended in the manual as an optimal cutoff score for 2- and 3-year-olds (Achenbach, 1992). The child's score on the CBCL/2-3 was classified as "Risk" if the score was on or above the cutoff point and "OK" if the score was below the cutoff point. On the ASQ:SE, for each potential cutoff point, the child's score was classified as "Risk" if the score was on or above the cutoff point being analyzed and "OK" if the score was below the cutoff point.

Using the CBCL/2-3 as a concurrent validity measure, the 24-month ASQ:SE yielded classification statistics that are shown in Tables 3 and 4. Three cutoff points (i.e., 1 standard deviations above the mean, 2 standard deviations above the mean, and the 3 semi-interquartile range) gave the same classification. In addition, two cutoff points (i.e., 1 standard deviation and the 2 semi-interquartile range) also gave identical classification results.

One cutoff point (i.e., the 1 semi-interquartile range) gave a third classification. Using the CBCL/2-3 as a concurrent measure, the 36-month questionnaire yielded classification statistics that are shown in Tables 5 and 6. As can be seen in the tables, it was found that three cutoff points (i.e., 1 standard deviations above the mean, 2 standard deviations above the mean, and the 3 semi-interquartile range) gave the same classification results. Again, it

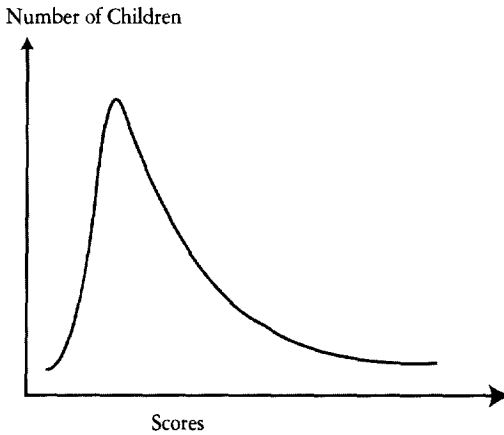
was also found that two cutoff points (i.e., 1 standard deviation and the 2 semi-interquartile range) gave identical classification results, and one cutoff point (i.e., the 1 semi-interquartile range) gave a third classification.

Step 5. The optimal cutoff point for each interval was determined by generating ROC curves based on the true positive and false positive scores shown in Tables 4 and 6.

That is, each point on the ROC curve was determined by plotting the true positive probability against the false positive probability for each possible cutoff point determined in Step 4.

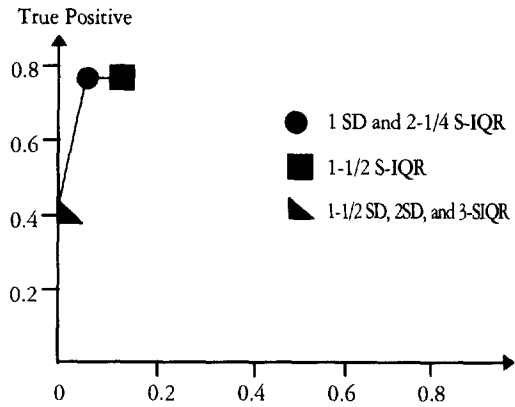
It was not possible to generate a ROC curve for the 24-month questionnaire because there were no true positive scores. However, the optimal cutoff score could be determined for the 24-month questionnaire by examining the specificity, false positive, over-referral, and under-referral scores. Three cutoff points (i.e., 1 and 2 standard deviations above the mean and the 3 semi-interquartile ranges) had the highest score for specificity at .98. The same cutoff points had the lowest score for false positive and over-referral. Therefore these three cutoff points were selected as the optimal cutoff points for the 24-month questionnaire. Even though the classification results were the same at 1 and 2 standard deviations and the 3 semi-interquartile range, the 3 semi-interquartile range was selected as the optimal cutoff point because of the data distribution. The cutoff score for the 3 semi-interquartile range (i.e., 77.5) was chosen over those of the standard deviations because the ASQ:SE scores did not

reflect a normal curve, making the standard deviations less useful.\* Figure 1 presents the curve of data distribution for the 24-month questionnaire.

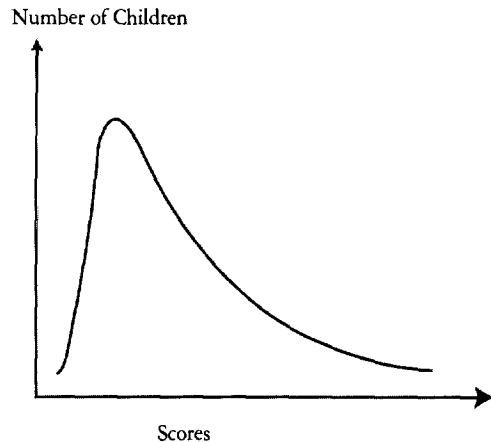


<Figure 1> Data distribution for the 24-month questionnaires

Three possible ROC curves for the 6-month questionnaire were plotted. These can be in Figure 2. From the ROC curve, two cutoff points that produced the same classification results (i.e., 1 standard deviation above the mean and the 2 semi-interquartile range) were chosen as optimal. Using the same rationale as used for the 24-month cutoff point, the 2<sup>1</sup>/<sub>4</sub> semi-interquartile range was selected as the optimal cutoff point for the 36-month questionnaire. Figure 3 presents the curve of data distribution for the 36-month questionnaire.



<FIGURE 2> ROC curve generated for the 36-month questionnaire using 1, 1½, and 2 standard deviations; and 1, 2¼, and 3 semi-interquartile ranges



<FIGURE 3> Data distribution for the 36-month questionnaire.

\*Semi-interquartile range is half the distance between the first and the third quartiles of a group of scores. When a distribution is not normal (i.e., skewed), interquartile range is more useful because interquartile range is not affected by outliers and less by non-normality.