

How Applicable Are Ages and Stages Questionnaires for Use With Turkish Children?

Emine Gul Kapci¹, Sevgi Kucuker², and Runa I. Uslu¹

Abstract

The majority of eligible children cannot access early intervention services in Turkey, often because they are not assessed. The authors adapted the *Ages and Stages Questionnaires* (ASQ) for Turkish children ages 3 to 72 months. Study participants consisted of 375 children who were classified as at risk for developmental delays, 564 children considered not at risk for developmental delays (both groups according to standardized assessments), and 39 children with known disabilities. The ROC analyses indicated that a two-domain criterion for classification would be used to classify children as at risk, unlike the original ASQ, which had a one-domain classification. The sensitivity, specificity, positive predictive value, and negative predictive value of ASQ were .94, .85, .97, and .75, respectively. Test–retest and interrater reliabilities calculated as overall percentage agreements were found to be 82% and 87%, respectively. The results demonstrated that the ASQ could be used to screen Turkish children who are at risk for developmental delays.

Keywords

children, infants, development, assessment, validity, reliability

Biological and psychosocial risk factors that affect children's development are more prevalent in developing countries (Ertem, 2004), and in these countries, more than 200 million children less than 5 years of age do not attain their developmental potential (Grantham-McGregor et al., 2007). According to the United Nations Children's Fund (2005), 3 million Turkish children have some form of disability and 16% of Turkish children ages birth to 5 years show inadequate development. In spite of the fact that preschool education is legally mandatory for children with developmental delays in Turkey, the Prime Minister's Office of Administration for Disabled People (2006) reported that only a minority of children with developmental delays who were less than 5 years old benefited from early childhood special education.

It has been well established that the early identification of and intervention for developmental problems are crucial to the developmental attainment of children. Early intervention programs for children with developmental problems have been shown to have lifelong benefits and also to be cost-effective in the long run. There are several reasons why potentially eligible children less than age 5 do not access early intervention services in Turkey. Perhaps the most important reason is the lack of a child-find system to identify children with developmental problems. At present, these children can only be identified if they are taken to a hospital or community health center for immunizations or as a result of some health problem. It would therefore not be erroneous

to assume that the state remains ignorant of the majority of young Turkish children with developmental problems for many years. In fact, research has shown that a majority of children with developmental problems are not identified before school age in many countries (Corrigan, Stewart, Scott, & Fee, 1996; Glascoe, 2000). The difficulties of conducting periodic screenings at early ages (e.g., limited number of professionals, time constraints, limits of child cooperation during health visit, attendance problems) may have contributed to the late identification of and provision of services for children with developmental problems.

In addition to a countrywide screening system, an early intervention service delivery system that takes into account a given country's resources is supposed to be implemented. For example, community health centers are well established all over Turkey, and the majority of parents in low- to middle-socioeconomic-status (SES) levels bring their children to these centers, mainly for immunizations. The utilization of parent-completed screening instruments in these community centers may help to identify many children who are in need

¹Ankara University, Turkey

²Pamukkale University, Denizli, Turkey

Corresponding Author:

Emine G. Kapci, Ankara University, Faculty of Educational Sciences, Cebeci, 06590 Ankara, Turkey
E-mail: kapci@education.ankara.edu.tr

Table 1. Summary of the Clinical Epidemiological Characteristics of the ASQ in Different Cultures

Study	Criterion measures	Age group (months)	Country/culture (n)	SE (%)	SP (%)	PPV (%)	NPV (%)
Squires et al. (1999)	<i>Bayley Scales of Infant Development-II, Gesell Development Schedule, Stanford-Binet Intelligence Scale, McCarthy Scales of Children's Abilities</i>	4-60	United States (1,648)	75.8	87.5	50.1	94.9
Skellern, Rogers, & O'Callaghan (2001)	<i>Griffith Mental Development Scales, Bayley Scales of Infant Development, McCarthy General Cognitive Intelligence Scale</i>	12, 18, 24, 48	Australia (136)	90	77	40	98
Kim & Sung (2007)	<i>Bayley Scales of Infant Development-II</i>	6-30	Korea (150)	50	78	24	92
Heo, Squires, & Yovanoff (2008)	<i>Korean-Denver II, Korean Development Indicators for the Assessment of Learning</i>	12-60	Korea (492)	81.2	91	-	-
Yu et al. (2007)	<i>Bayley Scales of Infant Development, Griffiths Tests, Denver-II</i>	12-60	19 countries (828)	87.4	82.3	13.3	99.5
Rydz et al. (2006)	<i>Battelle Developmental Inventory</i>	18	Canada (248)	67	39	34	71

Note: SE = sensitivity; SP = specificity; PPV = positive predictive value; NPV = negative predictive value.

of further assessment. In fact, utilization of parental screening instruments has increasingly become an adjunct to professionally administered tests, especially in the past decade. A primary advantage to using parent-completed instruments is that professionals can benefit from parents' extensive knowledge and experience regarding their own children. Another advantage is the opportunity to monitor a child's development regularly over time. Parental screening of child development is also cost-effective. In addition, research has demonstrated that parents can provide reliable and accurate information about their children's development regardless of their socioeconomic status, geographic location, or mental health (Coplan, 1982; Glascoe, 2003).

Whether completed by a parent or a professional, a good screening tool must have traditional psychometric properties (e.g., interrater and test-retest reliability, predictive and concurrent validity) as well as the clinical epidemiological characteristics of sensitivity, specificity, and positive and negative predictive values (Glascoe, 2005; Rydz, Shevell, Majnemer, & Oskoui, 2005). Although the *sensitivity* of a screening instrument is defined in terms of the proportion of children with disabilities that it identifies, *specificity* refers to the proportion of children without disabilities that it correctly identifies. Positive and negative predictive values are also considered to be important: A *positive predictive value* is the proportion of children with failing scores on a screening instrument who are later found to have a disability. The *negative predictive value* of a screening instrument is the proportion of children with passing scores on a screening instrument who are later found to be typically developing.

The set of *Ages and Stages Questionnaires* (ASQ) developed by Squires, Potter, and Bricker (1999) is considered to be a good parent-completion screening instrument that combines clinical epidemiological characteristics with traditional psychometric properties (Rydz et al., 2005). It has 19 age-specific questionnaires for assessing the development of children from 4 to 60 months in the domains of communication, gross motor, fine motor, problem solving, and personal-social development. Over time, a series of validity and reliability studies have been conducted, and cutoff scores have been calculated for these questionnaires from a database consisting of approximately 8,000 American children (Squires et al., 1999). According to the ASQ's guidelines (Squires et al., 1999), the development of a child is considered to be at risk when one or more of the domain scores is below the cutoff scores. The clinical epidemiological characteristics of the ASQ were calculated for 10 separate age groups consisting of 1,643 children and they were found to be favorable (see Table 1).

Overview of Studies on the ASQ's Adaptation to Other Cultures

The ASQ has been translated into Spanish, Danish, Korean, Norwegian, French and Chinese. A number of studies have examined the ASQ's traditional psychometric properties; several others have also investigated the ASQ's clinical epidemiological properties. In one of these studies, the mean domain scores of Norwegian children ($n = 1,341$; ages 4, 8, 12, 16, 20, 24, 30, 36, 48, and 60 months) were compared

with the U.S. normative data. Parental reports of children's development were found to be very similar in the two data sets (Janson & Squires, 2004). Another study with Norwegian children ($n = 1,172$; ages 4–60 months) reported that the child's ASQ scores significantly differed in relation to the mother's education, the child's gender, and whether the child was born prematurely or at full term. It was recommended that sex-specific norms should be developed, as there were differences in development between the sexes. These findings were reported to support the ASQ's construct validity (Richter & Janson, 2007).

Klamer, Lando, Pinborg, and Greisen (2005) compared the mean ASQ scores of extremely preterm and full-term Danish children, using 36- and 42-month questionnaires. Extremely preterm children's ASQ scores were found to be approximately 1 *SD* below the mean of the scores of the full-term children. The overall ASQ scores were found to be correlated significantly with IQ. The authors pointed out that the ASQ may serve as a useful instrument to assess developmental deficit in preterm children.

The ASQ were utilized in a study with 42 Korean infants ages birth to 12 months to examine the effects of maternal knowledge of infant development, maternal self-efficacy, and parenting behavior on child development (Seo, 2006). In this study, only Cronbach's alpha coefficients were calculated. The ASQ were shown to have a moderate internal consistency ($\alpha = .73$). Tsai, McClelland, Pratt, and Squires (2006) conducted a validity study with Taiwanese children age 36 months. This study demonstrated that the ASQ discriminated between children with ($n = 11$) and without developmental delays ($n = 101$). Results from parent and teacher ratings supported the interrater reliability of the ASQ.

A number of studies have examined the ASQ's clinical epidemiological characteristics (see Table 1). In one of these studies, Kim and Sung (2007) investigated ASQ's applicability to Korean children ages 6 to 30 months in the setting of a pediatric outpatient clinic. They found relatively low values for sensitivity and positive predictive value. By referring to the high negative predictive value, the authors concluded that the ASQ could be used as a screening tool. In another study with Korean children (Heo, Squires, & Yovanoff, 2008), the clinical epidemiological characteristics of the ASQ were examined in 11 age intervals with children ages 12 to 60 months. The researchers found low to good levels for sensitivity and specificity. Rydz et al. (2006) investigated whether the ASQ could be utilized in a busy pediatric clinic in Quebec, Canada. They found quite low levels for sensitivity, specificity, positive predictive value, and negative predictive value. The findings were partly attributed to the 3 months that lapsed between screening the child and performing the full assessment.

Skellern, Rogers, and O'Callaghan (2001) investigated the performance of the ASQ in an Australian sample of children who had been premature infants. Except for the positive

predictive value, they found high sensitivity, specificity, and negative predictive values for the ASQ. It was concluded that the ASQ could be used as a screening tool to detect developmental delays in these kinds of infants. In the same study, with a sample of children with known disabilities ($n = 21$) the ASQ's overall agreement with the criterion measure was found to be moderate (67%).

The effectiveness of magnesium sulphate in preventing eclampsia for women with pre-eclampsia were shown by the Magpie Trial study that was conducted in 33 countries with 10,141 women (Magpie Trial Collaborative Group, 2002). In a follow-up study conducted in 19 countries with a total of 2,046 children ages 12 to 60 months born to women in the trial, the ASQ's performance in predicting children with severe neurosensory disability was examined (Yu et al., 2007). Of these children, 828 had been assessed using both the ASQ and a neurosensory assessment. This study also supported the ASQ's applicability to diverse populations and its utilization in a variety of settings.

Research Questions

The lack of a screening system, insufficient number of professionals, time constraints, and the limited numbers of professionally administered screening tests are the prominent obstacles preventing professionals from identifying Turkish infants/children with developmental problems. Most important, there have been no parent-completed screening tools that would be an adjunct to professional assessment in identifying children who are at risk for or who have developmental delays.

The ASQ meets the standards of a good screening instrument, with both traditional psychometric properties and clinical epidemiological characteristics. Thus, in addition to assessing ASQ's validity and reliability, in this study we wanted to examine its sensitivity, specificity, and positive and negative predictive values for Turkish children. To our knowledge, the present study is the first to examine the clinical epidemiological characteristics of all 19 age-specific questionnaires of the ASQ in another culture. The research questions are as follows:

1. Is the Turkish ASQ (ASQ-TR) culturally appropriate?
2. What is the validity of the ASQ-TR?
3. What is the reliability of the ASQ-TR?

Method

Participants

From the cities of Ankara and Denizli, a total of 978 children within the 3- to 72-month age range participated in the study (see Table 2). Children were recruited from preschools,

Table 2. Children's Developmental Status Across Age Groups

ASQ month	Total (n)	Developmental status	
		Not at risk (n)	At risk/diagnosed (n)
4	41	28	13
6	39	26	13
8	41	29	12
10	46	30	16
12	44	28	16
14	34	24	10
16	32	22	10
18	33	21	12
20	27	17	10
22	30	13	16/1
24	38	25	13
27	41	22	18/1
30	56	23	29/4
33	34	23	11
36	43	20	18/5
42	81	40	33/8
48	100	57	38/5
54	54	22	28/4
60	164	94	59/11
Total	978	564	414

special education schools, hospitals, and community health centers. Of these children, 442 (45.19%) were girls and 536 (54.81%) were boys; 564 (67.89%) were "not at risk," whereas 375 (28.12%) were "at risk" for developmental delay. A total of 39 (4%) children in the sample had a clinical diagnosis for developmental disabilities within the 22- to 60-month age range (see Table 2). Of these children 16 (41%) were girls and 23 (59%) were boys. The diagnoses were as follows: Down syndrome ($n = 10$), autism ($n = 7$), cerebral palsy ($n = 4$), developmental delay ($n = 5$), pervasive developmental disorder not otherwise specified ($n = 4$), mental retardation ($n = 4$), speech and language disorder ($n = 4$), and visual impairment ($n = 1$).

In the study sample, a group of children were identified as prematurely born ($n = 165$, 16.9%). Following the suggestion by Squires et al. (1999), corrected ages were calculated for the premature children who were in the 4- to 24-month age range ($n = 67$, 6.85%).

The age of the mothers ranged from 18 years to 47 years ($M = 30.92$, $SD = 5.44$) and the majority were married (see Table 3). Approximately 40% of the mothers had a university education, and the remaining had elementary or high school educations. Slightly more than half of the mothers were unemployed.

As there was no available SES index for Turkish families, we made a categorization based on the Turkish Statistical Institute's (2006) poverty criteria, which were calculated based on the defining figure of approximately \$350. Accordingly, those families whose income was lower than twice this poverty level were categorized as low-income

Table 3. Demographics of the Mothers

Demographic	n	%
Education		
Elementary	276	28.9
High school	293	30.3
College/university	385	40.4
Marital status		
Married	929	95
Divorced	13	1.3
Separated/widowed	6	.6
No response	29	3
Employment		
Employed	409	41.8
Unemployed	540	55.2
No response	29	3
Family income		
Low	324	41.1
Middle	410	52
High	54	6.9

(<\$700); those families whose income ranged from 2 to 6 times the poverty level were categorized as middle-income (>\$700 and <\$2,100), and those families whose income was higher than 6 times the poverty level were categorized as high-income (>\$2,100). With respect to this categorization, the great majority of the mothers were from low- to middle-income families (Table 3).

Measures

Ages and Stages Questionnaires. The ASQ consists of 19 questionnaires relating to children of 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 27, 30, 33, 36, 42, 48, 54, and 60 months of age. The questionnaires have 30 developmental items divided into five domains: communication, gross motor, fine motor, problem solving, and personal-social development. Domain cutoff scores were calculated on the basis of 2 *SDs*. It was suggested that children whose questionnaire results stood at or below the established cutoff score in one or more of the domains should be referred for further assessment.

The questionnaire is completed by the parent or the caregiver of the child. Answers to the items are rated *yes*, *sometimes*, or *not yet* and are scored as 10, 5, or 0, respectively. Each different questionnaire is administered to the children of that age and in the ± 1 -month range. As the questionnaires cover wider age ranges after 27 months, administration of two successive forms has been recommended for children who are outside the ± 1 -month range (Squires et al., 1999).

To be able to include all the children outside of the ± 1 -month range, the age ranges of the questionnaires for children older than 36 months were expanded. Thus, the 42-month form included 39 to 44 months; the 48-month form included 45 to 50 months; the 54-month form included 51 to 56 months; and the 60-month form included 57 to 72 months. Children as old as 72 months may still be attending

preschools in Turkey; therefore, the 60-month questionnaire was used for these children. It should be noted that the rate of children who attend preschool is very low (33%; Turkish Republic of Ministry of Education, 2009) compared to rates for developed countries. Furthermore, the majority of these children—approximately 20% of 33%—attend preschools between the ages of 60 and 72 months. Therefore, utilization of the ASQ up to 72 months would be an important—if not the only—opportunity to identify those children who have or who are at risk of having developmental problems. A similar expansion of age ranges was applied to the 36-month form in an adaptation of the ASQ to Taiwanese culture (Tsai et al., 2006).

Translation and cultural appropriateness of the ASQ. The Turkish translation of the ASQ (ASQ-TR) was performed by the first and second authors of this article and its back-translation was done by the third author. The translated ASQs are referred to as the ASQ-TR. The cultural appropriateness of the ASQ-TR was scrutinized to ensure conceptual and experiential equivalence (Guillemin, Bombardier, & Beaton, 1993). Several professionals in the field of child development, special education, child psychiatry, and speech and language disorders ($n = 5$) were consulted about the appropriateness and wording of the items. Also, a group of eight mothers and six preschool teachers were asked to read the items and say whether there was any item that would be difficult for them to observe, understand, or fill in. All mothers had a primary education, and they were from low-income families. In line with the resulting suggestions, the following alterations were made: First, names of materials that not all Turkish children were familiar with, such as “Cheerios” or “stuffed animals,” were replaced with names of similar materials that all Turkish children would know. Second, certain behaviors that would be uncommon in Turkish cultural practice were replaced by more appropriate ones, for example, “Put the shoe on the table” was changed to “Put the book on the table.” Also it was observed that Turkish children generally give only their first name in answer to the question “What is your name.” Therefore, in addition to the original item (When you ask, “What is your name?” does your child say both her first and last names?), another item was added (When you ask, “What is your name and surname?” does your child say both her first and last names?). Responses to both questions were statistically analyzed to find out which item was more appropriate. Finally, a number of alterations were made in the communication domain by consulting an academician in the field of speech and language development and disorders in Turkish children. Accordingly, the original items were maintained, including the 14-month questionnaires. For the rest of the age-specific questionnaires, some modifications were performed: (a) either a new item was added or an existing item was moved to an earlier-age questionnaire and (b) some items were modified to better adhere to the structure of the Turkish language (see appendix for details).

Criterion measures. In the present study, one of three professionally administered instruments was used as a criterion measure:

- *Denver Developmental Screening Test* (2nd ed.; DENVER II; Frankenburg & Dodds, 1990; adapted by Anlar & Yalaz, 1995),
- *Gazi Early Childhood Development Assessment Scale* (GECNAS; Temel, Ersoy, Avci, & Turla, 2005), and
- *Ankara Development Screening Inventory* (ADSI; Savasir, Sezgin, & Erol, 1993).

The DENVER II was developed in the United States and adapted for Turkish children. The latter two instruments were developed in Turkey. GECNAS is a relatively new instrument, but the DENVER II and ADSI have been widely used in clinical and educational settings.

Although it is not recommended to utilize a screening test for validation of a new screening test, these are the only existing instruments available to assess children’s development in Turkey. As they have been shown to be valid and reliable instruments, they were utilized as a criterion measure in this study.

DENVER-II. The DENVER-II is a well-known test preferred for its practical use, and screens development in the birth to 6 age range. Its administration and scoring take about 20 to 30 min and require a trained interviewer. It consists of 125 items assessing four domains of development: fine motor and adaptation, gross motor, personal and social skills, and language skills. A given child’s test performance is categorized as questionable, abnormal, normal, or untestable. The Turkish standardization of the DENVER-II was completed by Anlar and Yalaz (1995). The interrater and test–retest reliabilities of the Turkish form were found to be 90% and 86%, respectively.

GECNAS. The GECNAS evaluates four domains of development: psychomotor, cognitive, language, and social–emotional in children ages birth to 72 months. It consists of 249 items and 21 age-specific sections. Each section has a different number of items (e.g., the section for 1-month-old children has nine items; for 61- to 72-month-old children, there are 19 items). Instead of a total score, the raw scores in each of the four developmental domains are placed on a graph and then compared to norms. A total of 4,242 children participated in a normative study of this instrument. Data from 1,890 children were used to test its split-half reliability ($r = .99$). The interrater reliabilities of the age groups were found to range from .88 to .99. Correlations of the subscale scores with the overall development score varied from .81 to .98. In addition, DENVER-II was used for GECNAS’s criterion validity.

ADSI. The ADSI is an instrument that assesses development in the first 6 years of life (birth–72 months). It contains a total of 154 items that are eventually grouped into four

developmental domains. The ADSI provides a total score as well as language–cognitive, fine motor, gross motor, and social skills–self-care domain scores. For the three age groups of birth to 12 months, 13 to 44 months, and 45 to 72 months, its test–retest reliabilities were reported as .99, .98, and .88, respectively, and Cronbach’s alphas of .98, .97, and .88, respectively. To examine the ADSI’s validity, data were gathered from three groups of children. The mean developmental scores of the children who were from families with low SES and children who lived in institutional care units were significantly lower than those of children who were from families with middle to high SES. These findings were reported to support the ADSI’s validity.

Procedure

The study proposal was approved by the Ethical Committee of Ankara University, School of Medicine. Written consent was obtained from the parents. Mothers of children were recruited from preschools; special education schools; the Pediatrics Departments of Ankara, Hacettepe, and Pamukkale University hospitals; and community health centers under the Ministry of Health. Permission was granted by the relevant administrative units of the universities. Permission was also obtained from the Ministry of Health to recruit parents from community health centers.

Data were collected between May 2005 and June 2006. The trained research assistants visited the above-mentioned institutions to invite parents to participate in the study. In the hospitals and in the community health centers, a room was allocated to the research assistants to meet parents and their children. The research assistants again met the parents of children attending preschools or special education schools in a meeting room or a test room. All the respondents were mothers. After a brief explanation about the aim of the study, mothers were asked to participate in the study. They were told that they would be informed about their children’s development on request. From preschools, special education schools, community health centers, and hospitals, 95%, 98%, 83%, and 77% of the mothers, respectively, agreed to participate in the study. The mothers who were approached through hospitals and who declined to participate reasoned that they had no time, had to go back home, or were busy with other health-related procedures in the hospital. Of the mothers who were asked to participate through preschools, special education schools, and community health centers but who declined to participate, most said that they were too busy or that they had participated in other studies. As no demographic information was obtained from mothers who declined to participate, we could not infer whether there was a difference between mothers who agreed to participate and those who declined to participate.

All participating mothers completed an information form that concerned demographic information about the children, such as gender, date of birth, gestational age, and family (e.g.,

parent’s age and education, marital and employment status, monthly family income). This was followed by the ASQ-TR. Subsequent to the completion of ASQ-TR, the research assistants administered one of the standardized developmental assessment tools (DENVER-II, GECDAS, or ADSI) within a few days at preschools, community health centers, or hospitals.

A great majority of the mothers completed the ASQ-TR at home (98% of mothers who were met in preschool; 70% of mothers from community health centers, and 86% of mothers from special education schools). The mothers who were recruited through hospitals filled in the ASQ-TR in allocated rooms and, on the same day, the research assistant administered a standardized instrument to the child. Approximately 3% of the ASQ-TRs were returned unfinished, these were excluded from the present analyses. Analyses were completed with the remaining data obtained from 978 participants.

Of these, 978 completed ASQ-TRs, 833 and 68 were used in the validity (receiver operating characteristic [ROC] and the concurrent validity) analyses and test–retest analyses, respectively. The remaining 77 ASQ-TRs were included in the interrater reliability and validity with known group analyses. The data used for the communication domain analyses were gathered from the children included in the ROC analyses and ages 16 months and older.

Results

In this section, findings of the modifications about the ASQ-TR’s communication domain are presented. Then, the ROC analyses and concurrent validity findings are reported. Finally, internal consistency and test–retest and interrater reliability results are presented.

Findings for the ASQ-TR Communication Domain

The appropriateness of the new items added to the communication domain or of the existing items that were moved to earlier age-specific questionnaires was examined. This analysis is based on the data gathered from children ages 16 months and older. These children were classified as not at risk ($n = 400$; 186 girls, 46.5%; 214 boys, 53.5%) or at risk ($n = 202$; 80 girls, 39.6%; 122 boys, 60.4%) according to standardized instruments. The correlations of the items with communication total scores for each relevant age group and their contribution to the internal consistency coefficient were calculated. In addition, a t test was conducted to compare the mean scores of the two groups of children.

The findings demonstrated that the modifications (i.e., the items that were added or moved to the early age-specific questionnaires or adhered to the structure of Turkish language) were by and large supported (see the appendix for details). Exceptions to these were the items that were added to the 16-, 22-, and 33-month questionnaires in which the two groups of children were found not to differ significantly.

However the first two items (i.e., the items added to the 16- and 22-month questionnaires) revealed high item-total correlations, and the item added to the 33-month questionnaire had a moderate item-total correlation. Nevertheless the latter was also included in the 33-month communication domain because of its considerable contribution to the overall Cronbach's alpha. In the light of these findings, the communication domain of the ASQ-TR consisted of seven items ranging from 16 months to 60 months.

Determination of ASQ-TR Cutoff Scores

To determine the optimal cutoff points of the ASQ-TR for each of the 19 age-specific questionnaires, ROC analyses (Gleitman, 1986) were performed. The ROC analyses strike a balance between sensitivity (proportion of true positives) and specificity (proportion of true negatives) by comparing true-positive probabilities against false-positive probabilities for a range of cutoff scores.

Prior to determination of the cutoff scores, the mean scores of boys and girls with typical development were compared for the five developmental domains. Analyses revealed that children's development did not differ with regard to gender, with two exceptions: girls had significantly higher mean scores than boys in the 24-month personal-social developmental domain, $t(23) = 2.10, p < .04$, and the 42-month communication domain, $t(39) = 3.56, p < .001$. Thus, data gathered from girls and boys were analyzed together.

In the present study, 833 children participated in the ROC analyses. To perform ROC analyses, each child was classified as at risk or not at risk according to criterion measures (GECEDAS, DENVER II, or ASDI). Of the 833 children, 565 were classified as not at risk and 268 were classified as at risk for developmental delays. ASQ-TR's classifications of children for a range of cutoff scores were then compared with the criterion measures' classification. The cutoff score that provided the optimal clinical epidemiological characteristics (highest sensitivity, specificity, PPV, NPV) were chosen for each domain.

As in the original study, children in the ASQ-TR were classified as developmentally at risk if at least one domain score was lower than cutoffs and as typically developing if all domain scores were above cutoffs. According to this classification, the overall percentage agreement of the ASQ-TR with the criterion measure was 71%. The overall sensitivity value of the ASQ-TR was quite high, meaning that 99.8 of every 100 children who required further evaluation were classified correctly. On the other hand, the overall specificity value indicated that only 57.5 of 100 children who did not require further assessment were classified correctly. Although the overall negative predictive value was high (99.39%), overall positive predictive value was somewhat lower (52.57%).

Acceptable values have been suggested to be at least 70% to 80% for sensitivity, 80% and above for specificity (Glascoe,

Table 4. Test Characteristics of the Overall ASQ-TR Based on One-Versus Two-Domain Criterion

ASQ-TR classification	Criterion measure classification		
	At risk	Not at risk	Total (n)
One domain^a			
At risk	266	240	506
Not at risk	2	325	327
Total	268	565	833
Two domains^b			
At risk	252	82	334
Not at risk	16	483	499
Total	268	565	833

Note: ASQ-TR = Turkish version of the *Ages and Stages Questionnaires*.

a. One domain: sensitivity = 99.25%; specificity = 57.52%; underreferral = 0.24%; overreferral = 28.81%; positive predictive value = 52.56%; negative predictive value = 99.38%; percentage agreement = 70.95%.

b. Two domains: sensitivity = 94.02%; specificity = 85.48%; underreferral = 1.92%; overreferral = 9.84%; positive predictive value = 75.45%; negative predictive value = 96.79%; percentage agreement = 88.24%.

2005), and 50% and above for positive predictive value (Rydz et al., 2005). Our findings indicated that specificity in particular was lower than the preferred values. Therefore, unlike that in the original ASQ, children with at least two domain scores below the cutoffs were classified as at-risk in the ASQ-TR, and its clinical epidemiological characteristics were reanalyzed. The findings are presented in Table 4.

With the two-domain criterion for classification, sensitivity and negative predictive values remained almost the same, whereas specificity and positive predictive values increased considerably in meeting the preferred levels. Not surprisingly, these values were also found to be better across all 19 age groups (see Table 5).

These results indicated that classification based on two domains below cutoffs seemed to be more appropriate. Therefore, Turkish children who received scores below cutoffs in at least two developmental domains would need to be referred for further assessment and children who received scores below the cutoff in one domain should be reassessed in due course.

Concurrent Validity

To evaluate concurrent validity, the percentage agreements between the ASQ-TR's classifications and those of the criterion measures were calculated. The overall percentage agreement of all age forms when classification was based on one domain was 70.95%, which increased to 88.24% when classification was based on two domains. Percentage agreements of classifications based on two domains were higher than 80% for all age groups except for the 4-month (67.50%) and 6-month (72.97%) forms of the ASQ-TR (see Table 5). On the whole, results from the 19 separate age forms of the ASQ-TR seem to confirm its concurrent validity.

Table 5. Clinical Epidemiological Characteristics of the ASQ-TR Across 19 Age-specific Questionnaires Based on One-Versus Two-Domain Criterion

Questionnaire	Domain criterion	SE (%)	SP (%)	UR (%)	OR (%)	PPV (%)	NPV (%)	PA (%)
4-month	One	100	39.28	0	42.50	41.38	100	57.50
	Two	75	64.28	7.5	25	47.37	85.71	67.50
6-month	One	100	23.07	0	54.05	35.48	100	45.94
	Two	100	61.54	0	27.02	52.38	100	72.97
8-month	One	100	55.17	0	32.50	45.83	100	67.50
	Two	90.91	79.31	2.50	15	62.50	95.83	82.50
10-month	One	100	53.33	0	32.55	48.15	100	67.44
	Two	92.31	90	2.32	6.98	80	96.43	90.70
12-month	One	100	60.71	0	28.20	50	100	71.79
	Two	90.91	82.14	2.56	12.82	66.66	95.83	84.61
14-month	One	100	75	0	18.75	57.14	100	81.25
	Two	87.50	87.50	3.12	9.37	70	95.45	87.50
16-month	One	100	54.54	0	32.26	47.37	100	67.74
	Two	100	95.45	0	3.22	90	100	96.77
18-month	One	90.91	61.90	3.12	25	55.55	92.86	71.87
	Two	81.82	85.71	6.25	9.37	75	90	84.37
20-month	One	100	47.06	0	36	47.06	100	64
	Two	100	76.47	0	16	66.66	100	84
22-month	One	100	28.57	0	35.71	58.33	100	64.28
	Two	78.57	92.86	10.71	3.57	91.67	81.25	85.71
24-month	One	100	76	0	18.75	53.85	100	81.25
	Two	85.71	84	3.12	12.50	60	95.45	84.37
27-month	One	100	52.17	0	28.95	57.69	100	71.05
	Two	100	95.65	0	2.63	93.75	100	97.37
30-month	One	100	44.44	0	30.61	59.46	100	69.39
	Two	100	92.59	0	4.08	91.67	100	95.92
33-month	One	100	60.87	0	30	43.75	100	70
	Two	100	82.61	0	13.33	63.63	100	86.67
36-month	One	100	80	0	13.51	70.59	100	86.49
	Two	91.67	96	2.70	2.70	91.67	96	94.59
42-month	One	95.45	56.10	1.59	28.57	53.85	95.83	69.84
	Two	95.45	85.36	1.59	9.52	77.78	97.22	88.89
48-month	One	100	58.82	0	29.17	50	100	70.83
	Two	100	80.39	0	13.89	67.74	100	86.11
54-month	One	100	61.54	0	22.73	64.28	100	77.27
	Two	94.44	96.15	2.27	2.27	94.44	96.15	95.45
60-month	One	100	67.06	0	23.14	56.25	100	76.86
	Two	100	91.76	0	5.78	83.72	100	94.21

Note: ASQ-TR = Turkish version of the *Ages and Stages Questionnaires*; SE = sensitivity; SP = specificity; PPV = positive predictive value; NPV = negative predictive value; PA = percentage agreement.

Validity With Known Groups

Validity for known groups was examined by the data obtained from a group of 39 children diagnosed as disabled. These children were recruited from special education schools (see the Participants section for details). When the classifications were based on one domain below the cutoff score, 35 (89.7%) of these children required further evaluation. On the other hand, when the classifications were based on two domains below the cutoff score, 34 (87.18%) of these children were found to require further evaluation.

Reliability

Internal consistency. To investigate the internal consistencies of the 19 age-specific questionnaires of the ASQ-TR, Cronbach's alpha, domain-total score correlations, and interdomain correlations were calculated. Cronbach's alpha ranged from .38 (4 months) to .95 (36 months) for communication, from .61 (6 months) to .93 (18 and 22 months) for gross motor, from .63 (16 months) to .93 (60 months) for fine motor, from .44 (20 months) to .93 (60 months) for problem solving, and finally from .50 (16 months) to .90 (60 months)

for personal–social development domains. The mean Cronbach's alphas for the full scale were higher for the older age groups than for the younger ones (e.g., 4 months $\alpha = .64$, 60 months $\alpha = .92$). Domain–total score correlations for each age-specific questionnaire of the ASQ-TR were all found to be statistically significant ($p < .05$). These correlations of domain scores with total scores were found to range between .54 and .95 for communication, .58 and .94 for gross motor, .71 and .96 for fine motor, .82 and .96 for problem solving, and .80 and .94 for personal–social development. The lowest and the highest interdomain correlations were found to be between the fine motor and communication domains ($r = .65$) and between the fine motor and problem-solving domains ($r = .79$), respectively. All the above results support ASQ-TR's internal consistency.

Test–retest reliability. The test–retest reliability was assessed by calculating the percentage agreement between classifications based on the questionnaires completed by mothers at 2-week intervals. Data were obtained from a total of 68 children from all age groups who were classified as typically developing ($n = 61$) and developmentally at risk ($n = 7$), according to criterion measures. These children were recruited from community health centers and preschools. The lowest and the highest percentage agreements were found to be .82 and .97 for the problem-solving and gross motor domains, respectively. The results revealed that the overall percentage agreements were slightly better for two-domain criteria than for one-domain criteria (.80 vs .82). The test–retest reliability findings showed that the ASQ-TR could reliably measure children's development.

Interrater reliability. To assess the interrater reliability of the ASQ-TR, the percentage agreement between children's classifications based on questionnaires completed by mothers and on questionnaires completed by teachers who knew the children for at least 3 months were calculated. These analyses were conducted separately for two groups of children. The first group of children ($n = 38$) were typically developing according to criterion measures and were attending preschools. The group consisted of 14 girls (36.8%) and 24 boys (63.2%) within the 42- to 60-month age range. The second group of children ($n = 39$) had known disabilities and were attending special education schools (see Participants section for details).

In the group of children with typical development, the overall percentage agreements between mothers and teachers were found to be 73.68% and 86.84% for classifications based on one and two domains, respectively. The overall percentage agreements, whether based on one- or two-domain cutoff scores, were identical (87.18%) for children with known disabilities.

Discussion

Timely identification of developmental problems and the provision of intervention in early childhood are highly critical as the beginning of a process that will lead to a variety of positive outcomes in children. The developmental assessment of

children in Turkey is carried out by professionals. As the number of professionals is quite low in Turkey, many children with developmental problems are not identified until they begin preschool or even elementary school; thus, they cannot benefit from early intervention services. The utilization of parent-completed instruments to identify developmental problems in young children therefore is crucial. In the present study, the ASQ, an instrument developed in the United States and also used in other countries, was investigated in terms of its applicability in the assessment of the development of Turkish children.

To test the ASQ's applicability, at first the conceptual and experiential equivalence were ensured for cultural appropriateness for Turkish children (Guillemin, Bombardier, & Beaton, 1993). This was done by taking into account the views of professionals, parents, and teachers. The appropriateness of the modifications made in the communication domain starting with the 16-month questionnaire was generally confirmed through statistical analyses. The findings for the communication domain in particular indicated that variations in language development as well as cultural context should be considered in the adaptation of developmental screening instruments to other cultures. Furthermore, every word, concept, or skill under scrutiny in the parent-completed developmental screening instruments may be important. For example, we had to not only replace the words *Cheerios* with *raisin* and *stuffed animal* with *furry toys* to meet the experiences of the Turkish mothers but also ensure that the items that included these terms should be conceptually equivalent in terms of the developmental behaviors under consideration.

In a study conducted by Heo et al. (2008), similar cultural modifications were made in several items (e.g., *Cheerios* was changed to *black bean*). The authors found that the differences between mean scores of the Korean ASQ data and U.S. ASQ data were especially apparent in the communication domain. Tsai et al. (2006) examined the cultural appropriateness of the ASQ for Taiwanese children by asking for the opinions of a group of experts, parents, and teachers. Only one expert raised a concern in relation to an image (i.e., about the use of a left hand holding a pair of scissors, because of a traditional preference to use right hands). In addition, the format was questioned by one parent (yes–no format) and one teacher (too many questions) instead of addressing cultural appropriateness. Referring to these comments, the authors stated that the 36-month ASQ items were culturally independent. The present study findings seem to support the idea that the ASQ items are culturally independent. However, there were some items that we had to alter into the Turkish culture and we also had to make several modifications to items of the communication domain to adapt them to the structure of the Turkish language.

Besides cultural appropriateness, findings of the validity and reliability analyses generally supported the applicability of the ASQ as a screening instrument for Turkish children. Perhaps the most salient findings were the difference between

the clinical epidemiological characteristics of the ASQ, based on one-domain criterion versus two-domain criterion, to identify children who are at risk or not at risk. As mentioned previously, Squires et al. (1999) classified children as at risk if their scores were below the specified cutoffs in one or more domain. Subsequent studies that reported the clinical epidemiological characteristics of the ASQ followed their one-domain criterion for classification (see Table 1). A close examination of these later studies revealed that they did not attain the preferred levels of sensitivity, specificity, or positive predictive values, except the ASQ study (Squires et al., 1999) conducted with American children.

In the present study, when a one-domain criterion was used to classify children who are at risk or not at risk, the specificity values in particular were found to be lower than the acceptable level in all of the 19 age groups (see Tables 4 and 5). Regarding the low levels of specificity, it is not surprising that we obtained a high level of overreferrals in many of the age groups. According to the specificity findings, nearly half of the Turkish children should be referred, when in fact they do not need further assessment. Especially in developing countries and Turkey as well, there is a limited number of professionals; consequently, the low rate of specificity and high rate of overreferrals indicate that the ASQ may have no value in practice for Turkish children. Nevertheless, when the two-domain criterion to classify children was used, quite high scores of specificity, positive predictive value, and low rates of overreferrals were obtained without much compromise from the overall sensitivity values. There was a trivial decrease in overall sensitivity (from 99% to 94%) but a considerable increase in specificity (from 57% to 85%) and positive predictive values (from 52% to 97%) reaching preferred levels.

The findings obtained from the analyses of the validity with known groups and the concurrent validity also supported the two-domain criterion. The percentage agreements were found to be 88% for the validity with known groups and 87% for concurrent validity in the present study, both of which were comparable to that of Squires et al. (1999), who found the former to be 96% and the latter to be 83%.

In relation to the reliability of the ASQ-TR, three types of analyses were conducted: The Cronbach's alpha coefficients were found to be generally favorable except in several age-specific forms (e.g., 4 and 6 months), and domain-total score correlations were also high, both of which supported the ASQ-TR's internal consistency. As far as we know, the ASQ's test-retest and interrater reliabilities, measured as percentage agreements, have not yet been investigated in other countries, apart from the original study in the United States (Squires et al., 1999). The test-retest (82%) and interrater reliability (87%) findings of the present study were somewhat lower than the U.S. study (both 94%). As a reliability of 80% or more was stated to be ideal for a screening instrument (Glascoe, 2005), it would not be erroneous to conclude that the ASQ-TR has the properties of both test-retest stability and interrater

reliability. It should be added that the reliabilities reported above were based on the two-domain criterion, which was considerably higher than the one-domain criterion, especially in the interrater reliability data gathered from children who are not at risk. These results provide further support for the suggestion that the two domain criterion for classification is more appropriate for the ASQ-TR.

Limitations of the Present Study

The present study has certain limitations. First, the numbers of children in few age groups were low (e.g., 20- and 22-month questionnaires). Nevertheless, there were children who were at risk and not at risk in each age-specific group, making it possible to calculate overreferral and underreferral rates. Needless to say, further studies with a larger sample size, especially in the above-mentioned age groups, are needed to confirm the present findings.

Another limitation was that the education level of the mothers was in general too high to represent the Turkish population; however, the relationship between education level and income is not linear in Turkey. As a result, we could state that the sample was more representative of Turkish population with regard to family income. Both Ankara and Denizli, from where the sample was recruited, are cities that receive a very large number of migrants every year. Most of the participants were recruited from community health centers to which families with lower income, in particular, are referred. Therefore the sample may be considered as representing a rural population as well as an urban one. Further studies with randomized countrywide samples are needed to confirm the findings of the present study.

A further limitation is related to the criterion measures. As there are no gold standard measures for Turkish children, a number of screening instruments were used for validation of the ASQ-TR.

It is important to note that the developmental progress can change for better or worse, especially in the early years (Glascoe, 2005); thus, repeated screening is essential to monitor children's development (American Academy of Pediatrics, 2001). In particular, children who fell below cutoffs in one domain in the ASQ-TR should also be taken into consideration and they should be monitored in successive assessments. Follow-up studies are necessary also to identify those children who are later found to have developmental problems and those children who are later found to outgrow their problems.

In summary, the results of the present study demonstrated that the ASQ-TR, as the first parent-completed developmental screening tool in Turkey, can be used to screen children with developmental problems. Moreover, the ASQ-TR is the first developmental screening instrument with the psychometric properties of sensitivity, specificity, and positive predictive values. We believe that the use of such parent-completed developmental tools would endorse attempts to set up a child-find system in Turkey.

Appendix

Modifications and Findings of the ASQ-TR Communication Domain

Item	ASQ months	Item-total correlation	Add to Cronbach's alpha	t test
Added item: "Does your child say words that indicate past time? For example: 's/he came' (<i>geldi</i>), 'it fell down' (<i>dustu</i>), 's/he has gone' (<i>gitti</i>), 's/he ate' (<i>yedi</i>). Please give an example_____."	16 18 20	.50 .70 .85	6 3 5	1.86 3.09** 2.88**
Added item: "Does your child make sentences such as 'I walked' (<i>yurudum</i>), 'I slept' (<i>uyudum</i>) or 'I played' (<i>oynadim</i>) to refer to past events? Ask your child questions such as 'What did you do in the park?' ('I slid') or 'What did you do with your friend?' ('I played'). Please give an example:_____."	22 24 27	.63 .64 .63	.1 .2 .4	1.20 2.66* 5.11**
Moved and adhered item: "Does your child use all of the words in a sentence (for example, 'a,' 'the,' 'am,' 'is,' 'are') to make complete sentences, such as 'I am going to the park' or 'Is there a toy to play with?' or 'Are you coming, too?'"	30 33 36	.74 .70 .73	1 6 1	7.58** 5.59** 4.69**
[This item, which is included in the 42-, 48-, and 54-month ASQ questionnaires, was moved to the 30-, 33-, and 36-month questionnaires in the ASQ-TR and modified to better adhere to the structure of the Turkish language, as follows: "Does your child use all of the words in a sentence (for example, 'e,' 'da,' 'mi') to make complete sentences, such as 'I am going home' ('Eve <i>gidiyorum</i>), or 'My sister/brother is at school' (' <i>Kardesim okulda</i> ') or 'Is there a chewing gum?' (' <i>Sakiz var mi?</i> ')?"]	33 36 42 42	.29 .51 .78 .79	7 4 3 2	1.72 4.34** 6.48** 7.55**
Added item: "When you ask, 'What is your name and surname?' does your child say both her first and last names?"	48	.72	2	6.59**
Moved and Adhered Item: "Does your child use endings of words, such as 's,' 'ed,' and 'ing'? For example, does your child say things like, 'I see two cats,' 'I am playing,' or 'I kicked the ball?'"	54 60	.58 .71	— —	4.98** 9.02**
[This item, which is included in the 48- and 54-month questionnaires of the ASQ, was moved to the 42-month questionnaire of the ASQ-TR and modified to better adhere to the structure of the Turkish language, as follows: "Does your child use words together with plurals or tenses? For example, 'Cats had gone' (<i>Kediler gitti</i>), 'I am playing' (<i>Oynuyorum</i>), or 'I threw the ball' (<i>Topu attim</i>)."]				
Added item: "Does your child make a complete sentence by using words such as 'and,' 'but,' or.' For example, 'I came home but my mum wasn't at home,' 'I ate food and drank water.' Please give an example_____."				

Note: In the original *Ages and Stages Questionnaires (ASQ)*, each item is repeated in two or three subsequent age-specific questionnaires. A similar approach was followed for the added/adhered items into communication domain.

* $p < .05$. ** $p < .01$.

Authors' Note

Preliminary findings were presented at the 18th World Congress of the International Association for the Child and Adolescent Psychiatry and Allied Professions (IACAPAP) as poster presentation, April 30 to May 3, 2008, Istanbul.

Acknowledgments

We express our gratitude to Pinar Ege, Professor of Speech and Language Development and Disorders, at Ankara University for her review and suggestions of the items in the communication domain of the ASQ-TR.

Declaration of Conflicting Interests

The author(s) declared no conflicts of interests with respect to the authorship and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research and/or authorship of this article: This research was supported by a grant from the Scientific and Technical Research Council of Turkey (Turkiye Bilimsel ve Teknik Arastirma Kurumu; TUBITAK) Project No.105K003.

References

- American Academy of Pediatrics. (2001). Developmental surveillance and screening of infants and young children. *Pediatrics, 108*, 192–196.
- Anlar, B., & Yalaz, K. (1995). *Denver gelismisel tarama testinin Turk cocuklarına uyarlanması ve standardizasyonu* [Denver developmental screening test's adaptation and standardization to Turkish children]. Ankara, Turkey: Meteksan AS.
- Coplan, J. (1982). Parents' estimate of child's developmental level in high-risk population. *American Journal of Diseases of Children, 136*, 101–104.
- Corrigan, N., Stewart, M., Scott, M., & Fee, F. (1996). Predictive value of preschool surveillance in detecting learning difficulties. *Archives of Disease in Childhood, 74*, 517–521.
- Ertem, I. O. (2004). Early childhood development in developing countries. In S. Parker, B. Zuckerman, & M. Augustyn (Eds.), *Developmental and behavioral pediatrics: A handbook for primary care* (2nd ed., pp. 75–78). Philadelphia, PA: Lippincott Williams & Wilkins.
- Frankenburg, W. K., & Dodds, J. B. (1990). *Denver II screening manual*. Denver, CO: Denver Developmental Materials.
- Glascoe, F. P. (2000). Detecting and addressing developmental and behavioral problems in primary care. *Pediatric Nursing, 26*, 251–257.
- Glascoe, F. P. (2003). Parents' evaluation of developmental status: How well do parents' concerns identify children with behavioral and emotional problems? *Clinical Pediatrics, 42*, 133–138.
- Glascoe, F. P. (2005). Screening for developmental and behavioral problems. *Mental Retardation and Developmental Disabilities, 11*, 173–179.
- Gleitman, H. (1986). *Psychology* (2nd ed.). New York, NY: Norton.
- Grantham-McGregor, S., Cheung, Y. B., Cueto, S., Glewwe, P., Rihter, L., & Strupp, B. (2007). Developmental potential in the first 5 years for children in developing countries. *Lancet, 369*, 60–70.
- Guillemin, F., Bombardier, C., & Beaton, D. (1993). Cross-cultural adaptation of health-related quality of life measures: Literature review and proposed guidelines. *Journal of Clinical Epidemiology, 46*, 1417–1432.
- Heo, K. H., Squires, J., & Yovanoff, P. (2008). Cross-cultural adaptation of a pre-school screening instrument: Comparison of Korean and U.S. populations. *Journal of Intellectual Disability Research, 52*, 195–206.
- Janson, H., & Squires, J. (2004). Parent-completed developmental screening in a Norwegian population sample: A comparison with US normative data. *Acta Paediatrica, 93*, 1525–1529.
- Kim, E. Y., & Sung, I. K. (2007). The Ages and Stages Questionnaire: Screening for developmental delay in the setting of a pediatric outpatient clinic. *Korean Journal of Pediatrics, 50*, 1061–1066.
- Klamer, A., Lando, A., Pinborg, A., & Greisen, G. (2005). Ages and Stages Questionnaire used to measure cognitive deficit in children born extremely preterm. *Acta Paediatrica, 94*, 1327–1329.
- Magpie Trial Collaborative Group. (2002). Do women with pre-eclampsia, and their babies, benefit from magnesium sulphate? The Magpie Trial: A randomised placebo-controlled trial. *Lancet, 359*(9321), 1877–1890.
- Republic of Turkey, Prime Ministry Administration for Disabled People. (2006). *Ozurlu cocuklara yonelik rehabilitasyon ve ozel egitim hizmetleri calismasi raporu* [The report of rehabilitation and special education services for children with disabilities]. Retrieved from <http://www.ozida.gov.tr>
- Richter, J., & Janson, H. (2007). A validation of the Norwegian version of the Ages and Stages Questionnaires. *Acta Paediatrica, 96*, 748–752.
- Rydz, D., Shevell, M. I., Majnemer, A., & Oskoui, M. (2005). Developmental screening. *Journal of Child Neurology, 20*, 4–21.
- Rydz, D., Srour, M., Oskoui, M., Marget, N., Shiller, M., Birnbaum, R., . . . Shevell, M. I. (2006). Screening for developmental delay in the setting of a community pediatric clinic: A prospective assessment of parent-report questionnaires. *Pediatrics, 118*, 1178–1186.
- Savasir, I., Sezgin, N., & Erol, N. (1993). *Ankara Gelisim Tarama Envanteri* [Ankara Developmental Screening Inventory] (2nd ed.). Ankara, Turkey: Turkish Psychological Association.
- Seo, S. J. (2006). A study of infant developmental outcome with a sample of Korean working mothers of infants in poverty: Implications for early intervention program. *Early Childhood Education Journal, 33*, 253–260.
- Skellern, C. Y., Rogers, Y., & O'Callaghan, M. J. (2001). A parent-completed developmental questionnaire: Follow up of preterm infants. *Journal of Pediatrics and Child Health, 37*, 125–129.

- Squires, J., Potter, L., & Bricker, D. (1999). *The ASQ user's guide* (2nd ed.). Baltimore, MD: Brookes.
- Temel, F., Ersoy, O., Avci, N., & Turla, A. (2005). *Gazi Erken Cocukluk Gelisimi Degerlendirme Araci* [Gazi Early Childhood Development Assessment Scale] Ankara, Turkey: Remay.
- Tsai, H. L. A., McClelland, M. M., Pratt, C., & Squires, J. (2006). Adaptation of the 36-month Ages and Stages Questionnaire in Taiwan: Results from a preliminary study. *Journal of Early Intervention, 28*, 213–225.
- Turkish Republic of Ministry of Education. (2009). T.C. Milli Egitim Bakanligi. Retrieved from <http://oocgm.meb.gov.tr/dosyalar.asp>
- Turkish Statistical Institute. (2006). *Başbakanlık Türkiye İstatistik Kurumu* [Ministry of Turkish Statistical Institute] Haber Bülteni. Sayı, 208. Retrieved from <http://www.tuik.gov.tr/PreHaberBultenleri.do?id=409>
- United Nations Children's Fund. (2005). *Erken cocukluk bakimi ve gelisimi* [Early childhood care and development] 2001–2005. Retrieved from http://www.unicef.org/turkey/hm/_hm3a.html
- Yu, L. M., Hey, E., Doyle, L. W., Farrell, B., Spark, P., Altman, D. G., Duley, L.; Magpie Trial Follow-Up Study Collaborative Group. (2007). Evaluation of the Ages and Stages Questionnaires in identifying children with neurosensory disability in the Magpie trial follow-up study. *Acta Paediatrica, 96*, 1803–1808.

About the Authors

Emine G. Kapci, PhD, is an associate professor of psychology at Ankara University. Her current interests include development of assessment instruments, emotional and behavioral problems of children.

Sevgi Kucuker, PhD, is an associate professor of special education at Pamukkale University. Her current interests include early intervention, parent education, children with mental retardation, siblings of children with special needs.

Runa I. Uslu, MD, is professor of child and adolescent psychiatry at Ankara University. Her current interests include infant mental health, child abuse, ADHD, and learning disorders.